THE KZB EQUATIONS ON RIEMANN SURFACES

GIOVANNI FELDER

Departement Mathematik, ETH-Zentrum,
CH-8092 Zürich, Switzerland

Abstract. In this paper, based on the author's lectures at the 1995 les Houches Summer school, explicit expressions for the Friedan–Shenker connection on the vector bundle of WZW conformal blocks on the moduli space of curves with tangent vectors at n marked points are given. The covariant derivatives are expressed in terms of "dynamical r-matrices", a notion borrowed from integrable systems. The case of marked points moving on a fixed Riemann surface is studied more closely. We prove a universal form of the (projective) flatness of the connection: the covariant derivatives commute as differential operators with coefficients in the universal enveloping algebra – not just when acting on conformal blocks.

1. Introduction

The Knizhnik–Zamolodchikov–Bernard, or KZB, equations are remarkable systems of differential equations which generalize the Gauss hypergeometric equation. There is such a system of differential equations for each simple complex Lie group G, a set of representations $V_1, \dots, V_n$ of G, an integer k and a non-negative integer h, with some restrictions. The equations are the equations of horizontality for a local section u of a vector bundle with connection, the vector bundle of conformal blocks, over the moduli space $\hat{M}_{h,n}$ of Riemann surfaces of genus h with tangent vectors at n marked points.

These equations have been studied mostly in the case of genus zero (the Knizhnik–Zamolodchikov equations), with a surprising range of applications, from number theory to quantum integrable systems to three-dimensional topology.

The genus one equations have also been studied in some detail. In particular, integral representations of solutions of hypergeometric type are known now both in genus zero and in genus one, see [33], [17].

For higher genus Riemann surfaces, the equations have not been studied in detail. One does know of the existence of the connection, and that it extends to a compactification of moduli space, leading to factorization theorems, but not much more is known. The object of this paper is a description in concrete terms of the KZB equations in genus ≥ 2. This is done in the original papers by Bernard [3, 4] for a particular parametrization of the moduli space of G-bundles based on a Schottky representation of the Riemann surfaces. In the case $n = 0$ a rather explicit and general description is given by Hitchin in [24]. Here we consider general coordinates on the moduli spaces, and relate the formulae to some familiar objects (r-matrices, ℓ-operators) of the theory of classical integrable systems. The setting is not quite the same as in integrable systems, so these objects come with a twist here and there.

The origin of the KZB equations is in the WZW model of conformal field theory [36, 30]. The central idea, formulated by Friedan and Shenker [16], is that the connection is given

by the energy-momentum tensor of the quantum field theory. The approach we take will be close to the original one of Friedan and Shenker, but we translate the notions to a more mathematical setting.

Here is an outline of the construction. Associated to the data $V_1, \dots, V_n$ and k large enough, we have a vector bundle over a suitable compactification of the moduli space of stable G-bundles on a Riemann surface Σ with marked points $p_1, \dots, p_n$.

A conformal block is a global holomorphic section of this vector bundle. The spaces of conformal blocks form a vector bundle over the moduli space $\hat{M}_{h,n}$. One then defines a "current insertion" $J_x(p)$ associated with a point p distinct from the marked points. It is a first order differential operator on conformal blocks corresponding to the vector field on the space of G-bundles defined roughly by the following local surgery procedure. If we have a G-bundle P, with given trivialization around p, and local coordinate t, the vector field points in the direction of the G-bundle obtained by taking out the fibers of P over a neighborhood of p and gluing them back with transition function $\exp(\epsilon x/(t-t(p)))$. The energy-momentum tensor $T(p)$ is essentially a suitable quadratic expression in the currents.

These operators depend on choices and do not map conformal blocks to conformal blocks. However, if u is a section of the vector bundle of conformal blocks then the covariant derivative $\nabla_\zeta u = \partial_\zeta u + \langle T, \zeta\rangle u$ is again a conformal block. Here $\zeta \in H^1(\Sigma, K^2 \otimes (-2\sum p_i))$ is a tangent vector that pairs with T, which depends on "flat" coordinates as a quadratic differential. See below for a more precise statement.

To make this explicit we choose to work in the double coset representation of the moduli space of G-bundle. We consider the case $G = SL(N, \mathbb{C})$ for simplicity of exposition. If U_S is a neighborhood of $S = \{p_1, \dots, p_n\}$, consisting of little disjoint open disks arount the points, isomorphism classes of G-bundles are in one-to-one correspondence with double cosets in $\mathcal{M}_G = G(U_S)\backslash G(U_S^\times)/G(\Sigma - S)$. Here $G(X)$ denotes the infinite-dimensional complex Lie group of holomorphic maps from the complex manifold X to G, and $U_S^\times = U_S - S$. The group $G(U_S^\times)$ has a central extension $\hat{G}(U_S^\times)$ by $\mathbb{C}^\times$ which splits over the two subgroups, and is a principal $\mathbb{C}^\times$-bundle over $G(U_S^\times)$. The space of conformal blocks, cf. [32], is the space of holomorphic functions $u(\hat{g})$ on $G(U_S^\times)$ with values in $V = V_1 \otimes \cdots \otimes V_n$ such that $u(\hat{g}z) = z^k u(\hat{g})$ for all $z \in \mathbb{C}^\times$ and $u(b\hat{g}n) = bu(\hat{g})$ for all $b \in G(U_S)$, $n \in G(\Sigma - S)$. The group $G(U_S)$ acts on the values by $bu = b(p_1) \otimes \cdots \otimes b(p_n)u$.

We will not be concerned here with the proper definitions of holomorphy in this infinite dimensional setting, and refer the interested reader to the papers [2, 13, 29]. The main result is that, with proper definitions, the space of conformal blocks is finite dimensional, and if we have any holomorphic family of Riemann surfaces with marked points, the spaces of conformal blocks form a holomorphic vector bundle over the parameter space.

The form of the connection is then, in terms of coordinates $\lambda_1, \dots, \lambda_n$ on the moduli space of G-bundles, and local trivialization

$$\nabla_\zeta u(\tau, \lambda) = \partial_\zeta u(\tau, \lambda) + \frac{1}{2\pi i} \oint_\gamma A(z)\zeta(z) dz\, u(\tau, \lambda),$$

for a second order differential operator $A(z)$ depending on coordinates like a quadratic differential (given a "flat structure" on Σ). Here γ is a contour winding around every point in S once, and a tangent vector ζ is represented by a holomorphic vector field on a pointed neighborhood $U_S^\times$ of S. The differential operator $A(z)$ is given by an expression

of the form

$$A(z) = \frac{1}{2(k+h^\vee)}(\mathrm{tr}^{(0)}\hat\ell(z)^2 + \mathrm{tr}^{(0)}q_1(z,\lambda)^{(0)}\hat\ell(z,\tau) + kq_2(z,\lambda))$$

The trace is the "trace over the auxiliary space" from integrable models: the object $\hat\ell(z)$ belongs to $\mathfrak{g}\otimes \mathrm{End}(V)\otimes\mathcal{D}$, where $\mathcal{D}$ denotes the space of differential operators in the variables λ_j, and $\mathrm{tr}^{(0)}$ is the trace over the first factor in the tensor product. The ℓ-operator $\hat\ell(z)$ is built out of the r-matrix by the formula

$$\hat\ell(z) = \sum_{i=1}^{m}\omega_i(z,\lambda)^{(0)}\partial_{\lambda_i} + \sum_{j=1}^{n} r(z,p_j,\lambda)^{(0j)} + kq(z,\lambda)^{(0)}.$$

In this expression $\omega_i(z,\lambda)dz$ is a basis of $H^0(\Sigma,\mathrm{Ad}(P_\lambda)\otimes K)$ Serre dual to the basis $\partial/\partial\lambda_j$ of the tangent space to the moduli space of G-bundle at the class of the G-bundle P_λ parametrized by λ. The "dynamical r-matrix" $r(z,t,\lambda)\in\mathfrak{g}\otimes\mathfrak{g}$ is the building block of all the other quantities. It is defined in 4.1. For instance, q_1 is the Lie bracket of the constant term in the Laurent expansion of r at its pole on the diagonal, and q_2 is obtained by applying the invariant bilinear form on the first order Laurent coefficient, see below.

2. Conformal blocks on Riemann surfaces

2.1. **Kac–Moody groups.** Let G be a simply connected complex simple Lie group, with Lie algebra $\mathfrak{g}$. For simplicity of exposition, we will assume that $G=\mathrm{SL}(N,\mathbb{C})$, the group of N by N matrices with complex entries and unit determinant.

For any complex manifold U, we let $G(U)$ be the group of holomorphic maps from U to G with pointwise multiplication. Is is an infinite dimensional complex Lie group. Its Lie algebra is the Lie algebra $\mathfrak{g}(U)$ of holomorphic maps $U\to\mathfrak{g}$.

Let U be an open disk containing the origin in the complex plane, and $U^\times = U - \{0\}$. The loop algebra $\mathfrak{g}(U^\times)$ has a universal central extension $\hat{\mathfrak{g}}(U^\times)$ (the affine Kac-Moody algebra) corresponding to the two-cocycle

$$c(x,y) = \mathrm{res}_{t=0}\mathrm{tr}(x'(t)y(t))dt \tag{1}$$

On the group level, see [31], we have a corresponding central extension

$$1\to\mathbb{C}^\times\to\hat G(U^\times)\xrightarrow{\pi} G(U)\to 1$$

which we now describe. For any smooth map $g:\Sigma\to G$ on a compact Riemann surface Σ with values in G, the WZW action [36] is ($\partial = dz\partial/\partial z$, $\bar\partial=\partial/\partial\bar z$ for any choice of local coordinate z)

$$W(g) = -\frac{i}{4\pi}\int_\Sigma \mathrm{tr}(g^{-1}\partial g\, g^{-1}\bar\partial g) + \frac{i}{12\pi}\int_B \mathrm{tr}((g^{-1}dg)^3)\in\mathbb{C}/2\pi i\,\mathbb{Z}.$$

In the second (Wess–Zumino) term, the Riemann surface Σ is viewed as the boundary of a three-dimensional manifold B and the map g is extended to a map from B to G. It can be shown that this is possible and that the integrals obtained by choosing different B's or different extensions differ by integer multiples of $2\pi i$. This follows from the fact that the difference of integrals corresponding to different choices is the integral over a closed three-manifold of the pull-back of the three-form $\frac{i}{12\pi}\mathrm{tr}((g^{-1}dg)^3)$ on G, which is in (the image of) $H^2(G,2\pi i\,\mathbb{Z})$. In the case of general simply connected G, one

replaces $\operatorname{tr}(xy)$ by (x,y), an invariant symmetric bilinear form, and $i(12\pi)^{-1}\operatorname{tr}((g^{-1}dg)^3)$ by $i(12\pi)^{-1}([g^{-1}dg,g^{-1}dg],g^{-1}dg)$, which is still a $2\pi i\times$integral three-form on G, provided the bilinear form is normalized in such a way that long roots have length squared two. In the non-simply connected case, things are slightly more tricky, see [15].

Let, for small $\epsilon>0$, $U_\epsilon=\{z\in U|\,|z|>\epsilon\}$ and let us first define the group $\hat G(U_\epsilon)$ consisting of equivalence classes of pairs (g,z) where g is a smooth map from U to G, holomorphic on U_ϵ and z is a non-zero complex number. Two pairs (g_1,z_1), (g_2,z_2) are equivalent if $g_2=g_1h$ for some smooth map $h:U\to G$ with $h|_{U_\epsilon}=1$ and

$$z_2=z_1e^{W(h)+\frac{1}{2\pi i}\int_U\omega(g_1,h)},\qquad \omega(g,h)=\operatorname{tr}(g^{-1}\partial g\bar\partial hh^{-1}).$$

Here $W(h)$ is defined by extending h to a smooth map from the Riemann sphere $\mathbb{P}^1=\mathbb{C}\cup\{\infty\}$ to G such that $h(z)=1$ if $|z|>\epsilon$. Let us denote by $[(g,z)]$ the equivalence class of (g,z).

The product in $\hat G(U_\epsilon)$ is defined by the rule

$$[(g_1,z_1)][(g_2,z_2)]=[(g_1g_2,z_1z_2\,e^{\frac{1}{2\pi i}\int_U\omega(g_1,g_2)})],$$

and the inverse of $[(g,z)]$ is $[(g^{-1},z^{-1}\exp(-\frac{1}{2\pi i}\omega(g,g^{-1})))]$. The WZW action obeys the product formula

$$W(gh)=W(g)+W(h)+\frac{1}{2\pi i}\int_\Sigma\omega(g,h)$$

for smooth maps $g,h:\Sigma\to G$. If $\Sigma=\mathbb{P}^1$, this formula can be used to check that the above relation is indeed an equivalence relation, and that the product is well defined on equivalence classes. Then, by construction, we have a central extension

$$1\to\mathbb{C}^\times\to\hat G(U_\epsilon^\times)\to G(U_\epsilon^\times)\to 1.$$

The nontrivial maps are $z\mapsto(1,z)$ and $(g,z)\mapsto g$. We may define $\hat G(U^\times)$ as the inverse image of $G(U^\times)\subset G(U_\epsilon)$. This is independent of the choice of ϵ up to isomorphism:

Proposition 2.1.

(i) *If $\epsilon_1<\epsilon_2$ are small, the restriction homomorphism $G(U_{\epsilon_2})\hookrightarrow G(U_{\epsilon_1})$ lifts uniquely to an embedding $j:\hat G(U_{\epsilon_2})\hookrightarrow\hat G(U_{\epsilon_1})$ preserving the central subgroup $\mathbb{C}^\times$. On the inverse image of $G(U^\times)$, j is an isomorphism.*

(ii) *If $U_1\subset U_2$ then the restriction homomorphism $G(U_2)\hookrightarrow G(U_1)$ lifts uniquely to an embedding $j:\hat G(U_2)\hookrightarrow\hat G(U_1)$ preserving the central subgroup $\mathbb{C}^\times$.*

The uniqueness follows from the fact that two lifts differ by a homomorphism from $G(X)$ to $\mathbb{C}^\times$, for appropriate X. Since $\mathbb{C}^\times$ is Abelian, any such homomorphism vanishes on the commutator subgroup. But $G(X)$ is equal to its commutator subgroup, see [31], Chapter 3.

Thus an element of $\hat G(U^\times)$ is an equivalence class of pairs (g,z) such that $g:U\to G$ is smooth and coincides with an an element of $G(U^\times)$ except on some small neighborhood of the origin.

By the property (ii), we can also pass to the inverse limit $\hat G=\lim_{U\ni0}\hat G(U^\times)$.

The central extension $\hat G(U^\times)$ splits over some subgroups of $G(U^\times)$:

Theorem 2.2.

(i) *The map $g\mapsto[(g,1)]$ is an injective homomorphism j of $G(U)$ into $\hat G(U^\times)$ with $\pi\circ j=\mathrm{Id}$.*

(ii) *Let U be embedded into a compact Riemann surface Σ, so that 0 is mapped to a point $p \in \Sigma$. Consider $G(\Sigma - \{p\})$ as the subgroup of $G(U^\times)$ of the maps extending to $\Sigma - \{p\}$. Then there is an injective homomorphism $j : G(\Sigma - \{p\}) \hookrightarrow \hat{G}(U^\times)$ with $\pi \circ j = \mathrm{Id}$. It is given explicitly by the formula*

$$g \mapsto [(\tilde{g}, e^{W(\tilde{g})})]$$

for any smooth map $\tilde{g} : \Sigma \to G$ which coincides with g on the complement of a small neighborhood of p.

We conclude this section by describing the Lie algebra $\mathrm{Lie}(\hat{G}(U^\times))$ of $\hat{G}(U^\times)$. First of all, the Lie algebra $\mathfrak{g}\,(U^\times)$ can be canonically embedded as a vector space into $\mathrm{Lie}(\hat{G}(U^\times))$: if $x \in \mathfrak{g}\,(U^\times)$ let x_{reg}, a regularization of x, be any smooth map $U \to \mathfrak{g}$ which coincides with x except possibly on some small neighborhood of the origin. Then we define the right-invariant vector field D_x acting on holomorphic functions on $\hat{G}(U^\times)$ by

$$D_x f(g) = \frac{d}{ds}\bigg|_{s=0} f([(\exp(-s x_{\mathrm{reg}}), 1)]g)$$

This definition is independent of the choice of regularization x_{reg}: the difference of the vector fields corresponding to different regularizations is $D_y f(g) = \frac{d}{ds}|_{s=0} f([(h_s, 1)]g)$ where $h_s = \exp sy$, and $y(z) = 0$ for $|z| > \epsilon$. But $(h_s, 1) \sim (1, \exp(-W(h_s)))$, so $D_y f = 0$ since $W(h_s) = O(s^2)$.

The Lie algebra $\mathrm{Lie}(\hat{G}(U^\times))$ is thus spanned by D_x, $x \in \mathfrak{g}\,(U^\times)$ and the generator K of the center: $Kf(g) = \frac{d}{ds}|_{s=0} f([(1, e^s)]g)$.

Proposition 2.3. *The Lie brackets on $\mathrm{Lie}(\hat{G}(U^\times))$ are given by the formulae (see (1))*

$$\begin{aligned} [D_x, K] &= 0, \\ [D_x, D_y] &= D_{[x,y]} + c(x,y)K, \qquad x, y \in \mathfrak{g}\,(U^\times). \end{aligned}$$

Thus $\mathrm{Lie}(\hat{G}(U^\times))$ is isomorphic to the affine algebra $\hat{g}(U^\times)$.

Proof: The first bracket follows from the fact that K is in the Lie algebra of a central subgroup. Let $g_t = \exp(-t x_{\mathrm{reg}})$ and $h_s = \exp(-s y_{\mathrm{reg}})$, for some choices of regularization. By definition,

$$D_x D_y f(g) = \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} f([(h_s, 1)][(g_t, 1)]g).$$

Neglecting terms of order higher than one in t or s, which do not contribute to the derivative at zero, we have

$$\begin{aligned} [(h_s, 1)][(g_t, 1)] &= [(h_s g_t, e^{\frac{1}{2\pi i}\int_U \omega(h_s, g_t)})] \\ &= [(g_t h_s \exp(-st[x,y]), e^{\frac{1}{2\pi i}\int_U \omega(h_s, g_t)})] \\ &= [(g_t, 1)][(h_s, 1)] \\ &\quad [(\exp(-st[x,y]), e^{\frac{1}{2\pi i}\int_U (\omega(h_s, g_t) - \omega(g_t, h_s))})]. \end{aligned}$$

Therefore,

$$\begin{aligned} D_x D_y f &= D_y D_x f + D_{[x,y]} f \\ &\quad + \frac{d}{dt}\bigg|_{t=0} \frac{d}{ds}\bigg|_{s=0} \frac{1}{2\pi i} \int_U (\omega(h_s, g_t) - \omega(g_t, h_s)) K f. \end{aligned}$$

The coefficient of Kf is

$$\frac{1}{2\pi i}\int_U \mathrm{tr}(dy_{\mathrm{reg}}dx_{\mathrm{reg}}) = \frac{1}{2\pi i}\oint \mathrm{tr}(y\,dx),$$

by Stokes' theorem. The contour of integration is chosen to lie in the region where x_{reg}, y_{reg} coincide with x and y. This integral is the residue at zero, which is what had to be proven. □

2.2. **Principal G-bundles.** Let Σ be a compact Riemann surface and $S = \{p_1, \dots, p_n\}$ be a finite set of $n \geq 1$ points on Σ. Let U_S be a neighborhood of S consisting of n embedded disjoint disks around the points p_i, and $U_S^\times = U_S - S$. Consider the group $G(U_S^\times)$ of holomorphic maps $U_S^\times \to G$. The group multiplication is defined pointwise. The group $G(U_S^\times)$ has the subgroups $G(U_S)$ of holomorphic maps from U_S to G, and $G(\Sigma - S)$ of holomorphic maps from $\Sigma - S$ to G.

For each element g of $G(U_S^\times)$ we have a principal G-bundle P_g on Σ: it is defined as the quotient space P_g of the disjoint union $U_S \times G \sqcup (\Sigma - S) \times G$, by the relation $U_S \times G \ni (z, h) \sim (z, g(z)^{-1}h) \in (\Sigma - S) \times G$, for all $z \in U_S^\times, h \in G$. The projection $P_g \to \Sigma$ is the projection onto the first factor, and the action of G is the right action on the second factor. Two G-bundles P_g, P_h are isomorphic if and only if there exist elements $b \in G(U_S)$ and $n \in G(\Sigma - S)$ such that $h(z) = b(z)g(z)n(z)$ for all $z \in U_S^\times$. Moreover it can be shown that every holomorphic G-bundle on Σ is isomorphic to some P_g. For $SL(N, \mathbb{C})$ this is the content of Grauert's theorem, see [23], Chapter 8. Thus equivalence classes of G-bundles are in one-to-one correspondence with double cosets in $M_S = G(U_S)\backslash G(U_S^\times)/G(\Sigma - S)$.

2.3. **Conformal blocks.** The group $G(U_S^\times)$ has a central extension $\hat{G}(U_S^\times)$ by $\mathbb{C}^\times$, which is a principal $\mathbb{C}^\times$-bundle over $G(U_S^\times)$. Let U_S be the union of disks U_i centered at p_i, and let $U_i^\times = U_i - \{p_i\}$. Then we have the direct product $\tilde{G}(U_S^\times) = \prod_{i=1}^n G(U_i^\times)$ of Kac-Moody groups, with its central subgroup $Z = (\mathbb{C}^\times)^n$. By definition $\hat{G}(U_S^\times)$ is the quotient of $\tilde{G}(U_S^\times)$ by the central subgroup

$$Z_0 = \{(z_1, \dots, z_n) \in Z \,|\, z_1 \cdots z_n = 1\}.$$

Alternatively, we may proceed as in Section 2.1 and define $\hat{G}(U_S^\times)$ directly to be the group of equivalence classes of pairs $[(g, z)]$, where g is a smooth map from U_S to G which coincides with a map in $G(U_S^\times)$ except on some small neighborhood of S. The equivalence relation is

$$(gh, ze^{W(h)+\frac{1}{2\pi i}\int_{U_S}\omega(g,h)}) \sim (g, z)$$

for all smooth maps $h : \Sigma \to G$, such that $h(q) = 1$ outside a small neighborhood of S.

The Lie algebra of $\hat{G}(U_S^\times)$ is spanned by vector fields D_x, $x \in \mathfrak{g}\,(U_S^\times)$ and the generator K of the center. The Lie brackets are as in Proposition 2.3, but the cocycle is now

$$c(x, y) = \sum_{j=1}^n \mathrm{res}_{p_j}\mathrm{tr}(dx\,y).$$

As before, the subgroups $G(U_S)$, $G(\Sigma - S)$ lift to subgroups of $\hat{G}(U_S^\times)$ which we denote be the same letters.

Let $V = V_1 \otimes \cdots \otimes V_n$ be a tensor product of finite-dimensional irreducible representations of G. Let $G(U_S)$ act on V by $gv = g(p_1) \otimes \cdots \otimes g(p_n)v$.

Definition: Let k be a non-negative integer. The space of conformal blocks $E_{V,k}(\Sigma, S)$ of level k is the vector space of holomorphic functions $u : \hat{G}(U_S^\times) \to V$ such that

$$u(bgn) = bu(g), \qquad u(gw) = w^k u(g),$$

for all $b \in G(U_S)$, $g \in \hat{G}(U_S^\times)$, $n \in G(\Sigma - S)$ and $w \in \mathbb{C}^\times$.

This definition is in fact independent of the choice of the union of disjoint disks U_S, in the sense that the restriction maps from the space of conformal blocks defined using U_S to the space defined using U'_S with $U'_S \supset U_S$ are isomorphisms.

Infinitesimally, the conditions on a conformal block u are as follows. Let γ by a contour winding each of the points p_j once.

Lemma 2.4. *Let $u \in E_{V,k}(\Sigma, S)$, and let $\hat{g} \in \hat{G}(U^\times)$ with $g = \pi(\hat{g}) \in G(U_S^\times)$. If $y \in \mathfrak{g}\,(U^\times)$ then*

$$D_y u(\hat{g}) = -\sum y(p_j)^{(j)} u(\hat{g}),$$

and if $\mathrm{Ad}(g^{-1})y$ extends to a map in $\mathfrak{g}\,(\Sigma - S)$, then

$$D_y u(\hat{g}) + \frac{k}{2\pi i}\oint_\gamma \mathrm{tr}(dg\, g^{-1}y)u(\hat{g}) = 0$$

Also, $Ku = ku$.

The first part is clear. As for the second, we have to go from the right action to the left action: In general if $x \in \mathfrak{g}\,(U_S^\times)$, let us define the right derivative

$$D^r_x u(\hat{g}) = \left.\frac{d}{ds}\right|_{s=0} u(\hat{g}\,[(\exp(sx_{\mathrm{reg}}), 1)]),$$

for any smooth x_{reg} on U_S equal to x away from a neighborhood of S. As for the left derivative, this object is independent of the choice of regularization.

Lemma 2.5. *For any $u \in E_{V,k}(\Sigma, S)$, $x \in \mathfrak{g}\,(U_S^\times)$ and $\hat{g} \in \hat{G}(U_S)^\times)$ with $\pi(\hat{g}) = g$,*

$$D^r_x u(\hat{g}) = -D_{\mathrm{Ad}(g_\lambda)x} u(\hat{g}) - \frac{k}{2\pi i}\oint_\gamma \mathrm{tr}(xg^{-1}dg)$$

Proof: Let $\hat{g} = [(g_{\mathrm{reg}}, w)]$, for some smooth g_{reg} on U_S. Let $h_{\mathrm{reg}} = \exp(sx_{\mathrm{reg}})$, where x_{reg} is a smooth map on U_S which coincides with x on the complement of a small neighborhood of S.

$$u([(g_{\mathrm{reg}}, w)][(h_{\mathrm{reg}}, 1)]) = u([(g_{\mathrm{reg}} h_{\mathrm{reg}} g_{\mathrm{reg}}^{-1}, 1)]\,\hat{g})e^{k\Gamma(g_{\mathrm{reg}}, h_{\mathrm{reg}})}, \tag{2}$$

with

$$\begin{aligned}\Gamma(g, h_{\mathrm{reg}}) &= \frac{1}{2\pi i}\int_{U_S} (\omega(g_{\mathrm{reg}}, h_{\mathrm{reg}}) - \omega(g_{\mathrm{reg}} h_{\mathrm{reg}} g_{\mathrm{reg}}^{-1}, g_{\mathrm{reg}})) \\ &= -s\frac{1}{2\pi i}\oint_\gamma \mathrm{tr}(g^{-1}dg\, x) + O(s^2).\end{aligned}$$

Now, $g_{\mathrm{reg}} x_{\mathrm{reg}} g_{\mathrm{reg}}^{-1}$ is smooth on U_S and coincides with $\mathrm{Ad}(g)x$ on the complement in U_S of a neighborhood of S. Taking the derivative at $s = 0$ of (2), we obtain the result. □

In particular this completes the proof of the second part of the preceding Lemma: take $x = \mathrm{Ad}(g^{-1})y$, and use the fact that, by right invariance, $D^r_x u = 0$, if $x \in \mathfrak{g}\,(\Sigma - S)$. The WZW action in Theorem 2.2, (ii) does not enter here, since it vanishes to second order in s if $h = \exp(sx)$.

The next important fact about conformal blocks is that if the representation associated to a point $p \in S$ is trivial, then the space of conformal blocks is canonically identified with the space associated to $S - \{p\}$. Consequently, we may think of conformal blocks as taking values in $\otimes_{x\in\Sigma} V_x$ where all but finitely many V_x are trivial.

Lemma 2.6. *Let $R \subset S$ be a non-empty subset and suppose $V_j = \mathbb{C}$, the trivial representation, if $p_j \notin R$. Let us embed $G(U_R^\times)$ into $G(U_S^\times)$ by extension by 1. This embedding lifts uniquely to an embedding $i : \hat{G}(U_R^\times) \hookrightarrow \hat{G}(U_S^\times)$, preserving the central subgroup $\mathbb{C}^\times$. Let $V' = \otimes_{p_j\in R} V_i$. Then the pull-back $i^* : u \mapsto u \circ i$ is an isomorphism from $E_{V,k}(\Sigma, S)$ onto $E_{V',k}(\Sigma, R)$.*

3. The connection

3.1. **The energy-momentum tensor.** The construction of the connection involves the introduction of differential operators associated to an additional point p which varies on the complement of S.

Let U_0 be some embedded disk in Σ, disjoint from U_S, and choose a local coordinate $z : U_0 \to \mathbb{C}$. Let $p \in U_0$, and set $U_{S\cup\{p\}} = U_S \cup U_0$. For any $x \in \mathfrak{g}$ let $x_j(p) \in \mathfrak{g}\,(U^\times_{S\cup\{p\}})$ be the map $q \mapsto x(z(q)-z(p))^j$, if $q \in U_0$ and $q \mapsto 0$ if $q \in U_S$. Denote by i_p^* the isomorphism $E_{V,k}(\Sigma, S \cup \{p\}) \to E_{V,k}(\Sigma, S)$ of Lemma 2.6. We define a differential operator $J_x(p)$, depending linearly on $x \in \mathfrak{g}$ and acting on conformal blocks:

$$J_x(p)u(\hat{g}) = D_{x_{-1}(p)} i_p^{*-1} u(i_p(\hat{g})), \qquad \hat{g} \in \hat{G}(U_S^\times).$$

This differential operator depends on the choice of local coordinate z. However the 1-form $J_x(p)dz(p)$ does not when acting on conformal blocks:

Lemma 3.1. *Let $J^z_x(p)$, $J^w_x(p)$ be the above differential operators defined using the local coordinates z and w respectively. Then for all p in the common domain of definition and $u \in E_{V,k}(\Sigma, S)$, $J^z_x(p)u(g) = (dw/dz)(z(p))J^w_x(p)u(g)$.*

Proof: We have

$$\frac{x}{z(q) - z(p)} = \frac{x}{w(q) - w(p)}\,\frac{dw}{dz}(z(p)) + \cdots\,,$$

where the dots stand for a function of q which is regular at p. This regular function does not contribute when acting on $E_{V,k}(\Sigma, S)$. Thus if $u \in E_{V,k}(\Sigma, S)$, then $J^z_x(p)u = J^w_{\tilde{x}}(p)u$ with $\tilde{x} = x(dw/dz)(z(p))$. The claim follows by linearity. □

Let us fix some local coordinate on U_0 and write simply $J_x(z)dz$ for $J_x(p)dz(p)$. Also, let $J(z)dz$ be the linear function on $\mathfrak{g}$ sending x to $J_x(z)dz$.

Proposition 3.2. *Let $u \in E_{V,k}(\Sigma, S)$ and $\hat{g} \in \hat{G}(U_S^\times)$ with $\pi(\hat{g}) = g$. Then*

(i) *The one-form $J(z)u(\hat{g})dz$ on U_0, with values in $\mathfrak{g}^* \otimes V$, has an analytic continuation to a one-form still denoted $J(z)u(\hat{g})dz$ on $\Sigma - S$, and $J(z)$ is a first order differential operator on conformal blocks. Moreover, for all $x \in \mathfrak{g}$, the differential operator*

$$J^S_x(z)u(\hat{g})dz = J_{\mathrm{Ad}(g(z)^{-1})x}(z)u(\hat{g})dz - k\,\mathrm{tr}(x dg(z)g(z)^{-1})u(\hat{g}),$$

defined on $U_S^\times$, extends to a meromorphic function of z with at most simple poles on S.

(ii) *If $y \in \mathfrak{g}\,(U_S^\times)$, then*

$$D_y^r u(\hat{g}) = -\frac{1}{2\pi i}\oint_\gamma \langle y(z), J(z)\rangle u(\hat{g})dz,$$
$$D_y u(\hat{g}) = \frac{1}{2\pi i}\oint_\gamma \langle y(z), J^S(z)\rangle u(\hat{g})dz.$$

The proof is deferred to 4.3, after we have introduced a more explicit description of how $J(z)$ acts on conformal blocks. At this point we only remark that by Lemma 2.5, the two formulae in (ii) are equivalent.

Now let $b_1,\dots,b_D$ be a basis of $\mathfrak{g}$ such that $\mathrm{tr}(b_ib_j)=\delta_{ij}$. Let $h^\vee = N$ be the dual Coxeter number of G. The *energy-momentum tensor* is the differential operator

$$T(p)u(g) = \frac{1}{2(k+h^\vee)}\sum_{j=1}^D D^2_{(b_j)_{-1}(p)} i_p^{*-1} u(i_p(g)).$$

Naively, $T(p)$ should transform as a quadratic differential under change of coordinate. However there is a correction term, due to the fact that after the first application of D_{b_j} the resulting function is not in general a conformal block. We proceed to give a formula for this correction term. If $w(z)$ is a holomorphic function of a complex variable, the *Schwarzian derivative* of w is the function

$$\{w,z\} = \frac{w'''(z)}{w'(z)} - \frac{3}{2}\left(\frac{w''(z)}{w'(z)}\right)^2.$$

Its main properties are the chain rule

$$\{w,z\}\left(\frac{dz}{du}\right)^2 + \{z,u\} = \{w\circ z, u\} \tag{3}$$

and the fact that $\{w,z\}=0$ if and only if $w(z)$ has the form $\frac{az+b}{cz+d}$, see, e.g., [35].

Lemma 3.3. *If $u \in E_{V,k}(\Sigma,S)$ and $T^z(p)$, $T^w(p)$ are the differential operators $T(p)$ defined using local coordinates z, w, respectively, then*

$$T^z(p)\,u = \left(\frac{dw}{dz}(z(p))\right)^2 T^w(p)\,u + \frac{1}{12}c_k\{w,z(p)\}\,u, \qquad c_k = \frac{kD}{k+h^\vee}$$

where $D=N^2-1$ is the complex dimension of G and $h^\vee = N$ is the dual Coxeter number of $\mathfrak{g}$.

Proof: Let $w(z)$ be the function expressing the coordinate w in terms of the coordinate z, and let us write z instead of $z(q)$ and $z_0 = z(p)$. Then we have the expansion in powers of $t = z - z_0$.

$$\frac{w'(z_0)}{w(z)-w(z_0)} = \frac{1}{t} - \frac{w''(z_0)}{2w'(z_0)} - \frac{1}{6}\{w,z_0\}t + O(t^2).$$

Correspondingly, for any $x\in\mathfrak{g}$, we have, indicating the choices of coordinate as superscripts,

$$w'(z_0)x^w_{-1}(p) = x^z_{-1}(p) - \frac{w''(z_0)}{2w'(z_0)}x^z_0(p) - \frac{1}{6}\{w,z_0\}x^z_1(p) + r_x$$

and r_x vanishes to second order at p. Now if u is a conformal block, then $D_{x_j(p)}u = 0$ for all $j \geq 0$, by invariance under $G(U_0) \subset \hat{G}(U^\times_{S\cup\{p\}})$. Thus we have (the derivatives are taken at z_0)

$$\begin{aligned} {w'}^2 D^2_{x^w_{-1}(p)}u(g) &= (D_{x^z_{-1}(p)} - \frac{w''}{2w'}D_{x^z_0(p)} - \frac{1}{6}\{w,z\}D_{x^z_1(p)})D_{x^z_{-1}(p)}u(g) \\ &= (D^2_{x^z_{-1}(p)} - \frac{w''}{2w'}[D_{x^z_0(p)}, D_{x^z_{-1}(p)}] \\ &\quad -\frac{1}{6}\{w,z\}[D_{x^z_1(p)}, D_{x^z_{-1}(p)}])u(g) \\ &= D^2_{x^z_{-1}(p)}u(g) - \frac{1}{6}\{w,z\}\, k\,\mathrm{tr}(x^2)u(g). \end{aligned}$$

By taking $x = b_j$, summing over j and dividing by $2(k+h^\vee)$ we get the desired result. □

We will need to know the behavior of $T(p)u$ under the action of $G(U_S)$ and $G(\Sigma - S)$.

Lemma 3.4. *For all $\hat{g} \in \hat{G}(U^\times_S)$,*

(i) $T(p)u(b\hat{g}) = b\,T(p)u(\hat{g})$, *if $b \in G(U_S)$.*
(ii) $\frac{d}{ds}\big|_{s=0}T(p)u(\hat{g}\exp(sx)) = J_{x'(p)}u(\hat{g})$, *if $x \in \mathfrak{g}\,(\Sigma - S)$*

Here $x'(p)$ denotes the derivative of x at p with respect to the local coordinate used to define T.

Proof: (i) is obvious. To prove (ii), we use the invariance of $i_p^{*-1}u$ under the right action of $G(\Sigma - (S\cup\{p\}))$: let $y(t) = x(t)$ for $t \in U_0$ and $y(t) = 0$ for $t \in U_S$. Then $i_p(\hat{g}\exp(sx)) = i_p(\hat{g})\exp(-sy)\exp(sx + O(s^2))$. Therefore, if $\kappa = k + h^\vee$,

$$\begin{aligned} 2\kappa\frac{d}{ds}\bigg|_{s=0} T(z)u(\hat{g}\exp(sx)) &= \frac{d}{ds}\bigg|_{s=0}\sum_{j=1}^{D} D^2_{(b_j)_{-1}(p)} i_p^{*-1}u(i_p(\hat{g}\exp(sx))) \\ &= \sum_{j=1}^{D} D_y D^2_{(b_j)_{-1}(p)} i_p^{*-1}u(i_p(\hat{g})). \end{aligned}$$

The last operator can be replaced by the commutator $[D_y, D^2_{(b_j)_{-1}(p)}]$, since $D_y i_p^{*-1}u = 0$ by the regularity of y on U_0. This commutator can be computed using Proposition 2.3. The four properties one has to use are: the invariance of $C = \sum b_j \otimes b_j$, i.e., $[C, x^{(1)} + x^{(2)}]$ for any $x \in \mathfrak{g}$; the fact that $h^\vee$ is half the Casimir of the adjoint representation, i.e., $\sum_j \mathrm{ad}(b_j)\mathrm{ad}(b_j) = 2h^\vee \mathrm{Id}_{\mathfrak{g}}$; the invariance of tr, i.e., $\mathrm{tr}([x,y]z) = \mathrm{tr}(x[y,z])$, in particular $\mathrm{tr}([x,b_j]b_j) = 0$; $D_y i_p^* u = 0$ if y is zero on U_S and regular on U_0. The result is

$$\sum [D_y, D^2_{(b_i)_{-1}(p)}] i_p^{*-1}u = D_a i_p^{*-1}u,$$

where $a(q) = x'(p)/(z(q) - z(p))$. □

3.2. **Flat structures.** We follow [35]. A *flat structure* on a compact Riemann surface Σ is an equivalence class of flat atlases $\{(U_\alpha, z_\alpha), \alpha \in I\}$. A flat atlas is covering of Σ by open sets U_α with local coordinates $z_\alpha : U_\alpha \to \mathbb{C}$ such that the transition functions $z_\alpha \circ z_\beta^{-1}$ are Möbius transformations. Two flat atlases are equivalent if their union is a flat atlas.

Flat structures exist on any Σ by the Riemann uniformization theorem. The description of all the flat structure on a compact Riemann surface Σ is given by the following result.

Theorem 3.5. *(See* [35]*) The set of flat structures on a compact Riemann surface Σ is an affine space over the vector space $H^0(\Sigma, K^2)$ of quadratic differentials.*

If we have two flat structures given by local coordinates z_α, w_α, respectively (we may assume that they are defined on the same covering by going to a refinement), the associated quadratic differential (their difference in the affine space) is given on U_α by the Schwarzian $\{z_\alpha, w_\alpha\}dw_\alpha^2$. The properties (3) of the Schwarzian derivative ensure that the quadratic differential is defined globally, independently of the choice of atlas within its equivalence class, and that this defines an affine structure on the set of flat structures.

Thus if we fix a flat structure and consider only coordinates in a flat atlas, we see that the energy momentum tensor depends on coordinates like a quadratic differential.

3.3. **Connections on bundles of projective spaces.** The Friedan–Schenker connection is a connection on the bundle of projective spaces of conformal blocks over the moduli space of curves (with additional data).

This means the following: if E is a holomorphic vector bundle over a complex manifold X, let $\mathbb{P}(E)$ be the bundle of projective spaces whose fiber at x is the projectivization $\mathbb{P}(E_x)$ of the fiber E_x. A connection on $\mathbb{P}(E)$ is an equivalence class of locally defined connections ∇_α on the holomorphic vector bundles $E|_{U_\alpha}$, for some open covering (U_α), such that on intersections $U_{\alpha,\beta} = U_\alpha \cap U_\beta$, $\nabla_\alpha - \nabla_\beta$ is a *scalar* holomorphic 1-form $a_{\alpha,\beta} \in H^0(U_{\alpha\beta}, T^*M)$. Two connections are equivalent if their difference is locally a scalar holomorphic one-form. Such data define a connection on $\mathbb{P}(E)$ globally: for each curve $t \mapsto \gamma(t)$ in M, the lift $\tilde\gamma(t) = \mathrm{cls}(s(\gamma(t)))$ with $\nabla_{\dot\gamma} s(\gamma(t)) = 0$, is uniquely and unambiguously determined by the initial condition $\tilde\gamma(0)$.

The curvature of a connection on $\mathbb{P}(E)$ is locally a two-form $F_\alpha = \nabla_\alpha^2$ with values in $\mathrm{End}(E)$ such that, on $U_{\alpha\beta}$, $F_\alpha - F_\beta = da_{\alpha\beta}$. Equivalent connections give rise to curvatures that differ locally by exact scalar one-forms.

A connection on $\mathbb{P}(E)$ is flat if F_α is a *scalar* (i.e., taking its values in the trivial subbundle of $\mathrm{End}(E)$ consisting of multiples of the identity) one-form. This means that contractible closed curve are lifted to closed curves in $\mathbb{P}(E)$.

There are two points of view when considering flat connections on $\mathbb{P}(E)$: the Čech and the Dolbeault point of view, in Hitchin's terminology [24]. The curvature of a flat connection is given by a set of closed two-forms F_α defined up to addition of exact forms, and such that $F_\alpha - F_\beta$ is exact on intersections. In the Čech description, we use the fact that locally F_α is exact to choose a representative ∇_α with $F_\alpha = 0$, so that a flat connection on $\mathbb{P}(E)$ is described by genuinely flat, locally defined connections on E. In the Dolbeault description, we may find smooth one-forms c_α such that $a_{\alpha\beta} = c_\alpha - c_\beta$. Then $\nabla_\alpha + c_\alpha$ is a globally defined connection on the vector bundle, but has non-trivial (scalar) curvature.

In our situation, the connection is described in the following terms. Let us fix the two-dimensional oriented smooth compact manifold Σ of genus h with marked points $S=\{p_1,\dots,p_n\}$. Fix also non-zero tangent vectors $v_1,\dots v_n$ at the points. Consider the moduli space $\hat M_{h,n}$ of complex structures I modulo the action of the group $\mathrm{Diff}_0(\Sigma,\mathrm{S},(\mathrm{v_j}))$ of orientation preserving diffeomeorphisms which leave the points and the tangent vectors fixed. This moduli space has a natural structure of algebraic variety of dimension $3(\mathrm{genus}(\Sigma)-1)+2n$, provided the genus is at least two, or one and $n\geq 1$, or zero and $n\geq 2$.

A tangent vector to $\hat M_{h,n}$ is a class in $H^1(\Sigma,T\Sigma\otimes O(-2S))$ by the Kodaira–Spencer deformation theory (S is identified with the divisor $p_1+\cdots+p_n$). If we choose local coordinates on a neighborhood of S, a class $[\zeta]$ is represented, in the Čech formulation, by a holomorphic vector field $\zeta(z)\frac{d}{dz}$ on $U_S^\times$, modulo vector fields extending to $\Sigma-S$ and vector fields extending to U_S and vanishing at S to first order[1]. If we view $I(p)$ as an endomorphism of the cotangent space at p with $I(p)^2=-1$, then an infinitesimal deformation $\dot I$ of I is described by a Beltrami differential $\mu=\mu(w)\frac{\partial}{\partial w}\otimes d\bar w$, whose local coordinate expression is defined by $\dot I\,dw=2i\mu(w)d\bar w$. The Beltrami differentials define the complex structure on the infinite dimensional manifold of complex structures on Σ: μ is the $(1,0)$ component of the tangent vector $\dot I$. The connection with the Čech description is that on U_S, $\mu=\bar\partial\zeta_S$ and on $\Sigma-S$, $\mu=\bar\partial\zeta_\infty$ for some smooth vector fields ζ_S, ζ_∞, defined up to addition of holomorphic vector fields such that ζ_S vanishes to first order at S. The difference $\zeta=\zeta_S-\zeta_\infty$ is holomorphic on $U_S^\times$.

The Friedan–Shenker connection is a connection on the projectivized vector bundle of conformal blocks over $\hat M_{h,n}$. Its fiber over $[I]$ is the projectivized vector space of conformal blocks $\mathbb{P}E_{V,k}(\Sigma,I,S)$, where we indicate the dependence on the complex strucure in the notation, for any choice of representative I. Conformal blocks associated to equivalent complex structures are canonically identified. A local holomorphic section of this bundle is represented by a holomorphic function $I\mapsto u(I,\hat g)$ invariant under the natural action of $\mathrm{Diff}_0(\Sigma,\mathrm{S},(\mathrm{v_j}))$, and such that $u(I,b\hat gn)=bu(I,\hat g)$ and $u(I,gz)=z^ku(I,g)$, see 2.3, for all $b(I,t)$ Given a local holomorphic section u, and a local holomorphic vector field ζ, we define the covariant derivative $\nabla_\zeta u$. To define the value of $\nabla_\zeta u$ at $\hat g$ in concrete terms, we need to specify how $\hat g$ changes as we deform the complex structure to take the derivative. For this we choose local coordinates to identify U_S with a fixed union of disks, so that an element $\hat g\in\hat G(U_S^\times)$ is defined even if we vary the complex structure. Thus we consider an enlarged moduli space of data (I,z), where I is again a complex structure and z is a local coordinate $z:U_S\to\sqcup_{j=1}^n\mathbb{C}$ defined on some neighborhood U_S, mapping p_j to the origin of the j^{th} copy of $\mathbb{C}$, and such that the tangent map $z_*(p_j)$ sends v_j to 1. The equivalence relation is given as before by the action of $\mathrm{Diff}_0(\Sigma,\mathrm{S},(\mathrm{v_j}))$. We have the natural projection $(I,z)\mapsto I$ from the enlarged moduli space to $\hat M_{h,n}$, with contractible fiber.

Let us now describe the tangent space at a point of the enlarged moduli space. If z is a complex coordinate for I on U_S we have by definition $I\,dz=i\,dz$. Differentiating, we see that a tangent vector $(\dot I,\dot z)$ at (I,z) is then given by a pair obeying $\dot I dz=i(1+iI)d\dot z$ or $\mu(z)=\frac{\partial}{\partial\bar z}\dot z$ on U_S. Thus a tangent vector $(\dot I,\dot z)$ gives rise to a vector field $\zeta=\zeta_S-\zeta_\infty$

[1]We say that a that a vector field vanishes to first order at a point p, if its components vanish and have vanishing first derivative, for some, and thus any, choice of coordinates at p. A vector field vanishes to first order at a set S if it vanishes to first order at all its points.

on $\mathfrak{g}(U_S^\times)$, with $\zeta_S = \dot z$, defined up to addition of a holomorphic vector field on $\Sigma - S$. Conversely, given a holomorphic vector field ζ on $U_S^\times$, we may write $\zeta = \zeta_S - \zeta_\infty$ as the difference of two smooth vector fields extending to U_S, $\Sigma - S$, respectively, and such that ζ_S vanishes to first order at S. This defines a tangent vector $(\dot I, \dot z)$ with $\mu = \bar\partial\zeta_S$ on U_S, $\mu = \bar\partial\zeta_\infty$ on $\Sigma - S$ and $\dot z = \zeta_S$. Different choice of ζ_S, ζ_∞ lead to tangent vectors differing by infinitesimal diffeomorphisms.

In other words, a tangent vector to the enlarged moduli space is the same as a holomorphic vector field $\zeta(z)\frac{d}{dz}$ on $U_S^\times$ defined modulo holomorphic vector fields extending to $\Sigma - S$.

Let $u(I, \hat g)$ be a local holomorphic section of the vector bundle of conformal blocks. Thus u is a holomorphic map $I \mapsto u(I, \cdot) \in E_{V,k}(\Sigma, I, S)$, invariant under the natural action of $\mathrm{Diff}_0(\Sigma, \mathrm{S}, (\mathrm{v_j}))$.

To define the covariant derivative in the direction of a tangent vector $[\zeta]$, we choose a one parameter family (I_τ, z_τ), $|\tau| < \epsilon$, be a holomorphic of data in the enlarged moduli space, so that the class of $\dot I = \frac{d}{d\tau}|_{\tau=0} I_\tau$ is $[\zeta]$. If $\hat g = [(g_{\mathrm{reg}}, 1)]$ is in $\hat G$ for the complex structure $I = I_0$, then $\hat g_\tau = [(g_{\mathrm{reg}} \circ z^{-1} \circ z_\tau, 1)]$ is in $\hat G$ for the complex structure I_τ (its expression in the local coordinate z_τ coincides with the expression of $\hat g$ in terms of the local coordinate z). Let ζ be a vector field corresponding to $(\dot I_\tau, \dot z_\tau)$ at $\tau = 0$.

The covariant derivative is defined by an expression

$$\nabla_\zeta u(I, \hat g) = \frac{d}{d\tau}\bigg|_{\tau=0} u(I_\tau, \hat g_\tau) + A(I, \zeta) u(I, \hat g), \qquad z = z_{\tau=0}. \tag{4}$$

Here $A(I, \zeta)$ is a differential operator on conformal blocks depending linearly on ζ. In order for this formula to define a connection on $\hat M_{h,n}$ we need to show that

$$A(I, \eta) = 0, \tag{5}$$

if η extends to a holomorphic vector field on $\Sigma - S$ and that if η is holomorphic on U_S and vanishes to first order at S, then

$$A(I, \eta)u(\hat g) = D_{\eta g' g^{-1}} u(\hat g) + \frac{k}{4\pi i}\oint_\gamma \eta(z)\mathrm{tr}((g'(z)g(z)^{-1})^2) dz\, u(\hat g) \tag{6}$$

Here $g = \pi(\hat g)$.

The origin of (6) is the following: if we replace z_τ by some other coordinate w_τ such that $w_0 = z_0$ and such that $\dot w_0 - \dot z_0 = \eta$, then the tangent vector ζ is replaced by $\zeta + \eta$. the right-hand side of (4) is replaced by

$$\frac{d}{d\tau}\bigg|_{\tau=0} u(I_\tau, [(g_{\mathrm{reg}} \circ z^{-1} \circ w_\tau)]) + A(I, \zeta + \eta) u(\hat g), \tag{7}$$

which is the same as (4), if we have (6), as a straightforward calculation shows.

Summarizing, we have:

Proposition 3.6. *Let for I in some open set of the space of complex structures on Σ, and ζ a holomorphic vector field defined on some pointed neighborhood of S, $A(I, \zeta)$ be a differential operator acting on $E_{V,k}(\Sigma, I, S)$, depending linearly on ζ. Then $A(I, \zeta)$ defines locally via (4) a connection on the bundle of projective spaces of conformal blocks, if (i) ∇_ζ maps local sections to local sections, (ii) $A(I, \zeta) = 0$ if ζ extends to a holomorphic vector field on $\Sigma - S$ and (iii) eq. (6) holds if η is regular at S and vanishes there to first order.*

3.4. **The Friedan–Shenker connection.** We now claim that the energy momentum tensor can be used to construct a differential operator A obeying the hypotheses of Proposition 3.6, and thus defines a connection on $M_{h,n}$.

Proposition 3.7. *Let $\kappa = k + h^\vee$. Let us fix a flat structure on (Σ, I). and let z be a coordinate on U_0. Then $T(p)dz(p)^2$, defined for $p \in U_0$, extends, as a function of p, to a differential-operator-valued quadratic differential still denoted $T(p)dz(p)^2$ on $\Sigma - S$. Moreover, if $p \in U_S^\times$,*

$$T^S(p)dz(p)^2u(\hat g) = \big(T(p)dz(p)^2 + \mathrm{tr}^{(0)}\big((g(p)^{-1}dg(p))^{(0)}J(p)\big)dz(p) + \frac{k}{2}\mathrm{tr}((g^{-1}(p)dg(p))^2)\big)u(\hat g),$$

in local coordinates on U_S, extends to a meromorphic quadratic differential on U_S with at most poles of second order on S. Here the square in the last term is the symmetric square, sending differentials to quadratic differentials.

The proof is deferred to 4.3.

Theorem 3.8. *Let $T(z)dz^2$ be a local coordinate expression, for some choice of flat structure, of the energy momentum tensor and $\zeta(z)\frac{d}{dz}$ a holomorphic vector field on $U_S^\times$. Then $A(I,\zeta) = \frac{1}{2\pi i}\oint_\gamma \zeta(z)T(z)dz$ defines locally, via (4), a connection on the vector bundle of conformal blocks.*

Proof:We show that the criteria of Proposition 3.6 are satisfied. We first show, using Lemma 3.4 and Proposition 3.2 (ii), that the connection is well-defined, i.e., the covariant derivative maps local sections to local sections. Clearly $\nabla_\zeta u(b\hat g) = b\nabla_\zeta u(\hat g)$ by Lemma 3.4 (i). Let $n \in G(\Sigma - S) = G(\Sigma - S, I)$, for the complex structure I, and suppose n is part of a holomorphic family $\tilde n_\tau$, such that $\tilde n_\tau \in G(\Sigma - S, I_\tau)$. We must compare this deformation with the deformation $n_\tau = n \circ z^{-1} \circ z_\tau$ appearing in the definition of the connection. We have, by the invariance of u,

$$\begin{aligned} \frac{d}{d\tau}\bigg|_{\tau=0} u(I_\tau, (\hat g n)_\tau) &= \frac{d}{d\tau}\bigg|_{\tau=0} u(I_\tau, \hat g_\tau n_\tau) \\ &= \frac{d}{d\tau}\bigg|_{\tau=0} u(I_\tau, \hat g_\tau n_\tau \tilde n_\tau^{-1}) \\ &= \frac{d}{d\tau}\bigg|_{\tau=0} u(I_\tau, \hat g_\tau) + D^r_{\zeta n' n^{-1}} u(I, \hat g). \end{aligned}$$

In the last term the prime means the derivative with respect to the local coordinate z on U_S. The point is that the fact that $\tilde n_\tau$ extends to $\Sigma - S$ implies that

$$\frac{\partial}{\partial\tau}\bigg|_{\tau=0} \tilde n_\tau + \zeta n'$$

extends to a holomorphic function on $\Sigma - S$ with values in $\mathfrak g$ for the complex structure I. This translates into the necessary condition for A to define a connection:

$$A(I,\zeta)u(\hat g n) = A(I,\zeta)u(\hat g) - D^r_{\zeta n' n^{-1}} u(I, \hat g),$$

or, infinitesimally ($G(\Sigma - S)$ is connected),

$$\frac{d}{ds}\bigg|_{s=0} A(I,\zeta)u(\hat g \exp(sx)) = -D^r_{\zeta x'} u(I, \hat g),$$

for all $x \in \mathfrak{g}\,(\Sigma - S)$. This property follows from Lemma 3.4 (ii) with Proposition 3.2.

The property (ii) of Proposition 3.6 follows from the fact stated in Proposition 3.7 that $T(z)dz^2$ is a holomorphic quadratic differential on $\Sigma - S$, by Stokes' theorem.

By the same argument, we have $\oint T^S(z)\eta(z)dz = 0$, if $\eta(z)d/dz$ is holomorphic on U_S. Thus, by the formula in Proposition 3.7, we obtain

$$\begin{aligned} 0 &= A(\eta) + \frac{1}{2\pi i}\oint_\gamma \langle (g^{-1}dg, J\rangle \eta + \frac{k}{4\pi i}\oint_\gamma \eta(z)\mathrm{tr}((g'g^{-1})^2)dz \\ &= A(\eta) + \frac{1}{2\pi i}\oint_\gamma \langle dgg^{-1}, J^S\rangle \eta - \frac{k}{4\pi i}\oint_\gamma \eta(z)\mathrm{tr}((g'g^{-1})^2)dz \\ &= A(\eta) - D_\eta g'g^{-1} - \frac{k}{4\pi i}\oint_\gamma \eta(z)\mathrm{tr}((g'g^{-1})^2)dz, \end{aligned}$$

which is (iii) of Proposition 3.6. □

Remark. The dependence on the choice of flat structure is given by Lemma 3.3. Thus if we change flat structure we get an equivalent connection in the sense of 3.3.

4. The Knizhnik–Zamolodchikov–Bernard equations

We give here formulae for the connection. We assume that the genus of Σ is at least two. The cases of genus zero and one require slight modifications.

4.1. **Dynamical r-matrices.** In this section, some of the results of [8] are reviewed. The setting is the same as in Section 2.2. Furthermore, we assume that the genus of Σ is at least two.

Proposition 4.1. *Let $g \in G(U^\times)$ be such that $H^0(\Sigma, \mathrm{Ad}(P_g)) = 0$. Let C be the invariant symmetric tensor $\sum_j b_j \otimes b_j$ for any basis basis (b_j) of $\mathfrak{g}$ so that $\mathrm{tr}(b_ib_j) = \delta_{ij}$. Let $\xi_1, \dots, \xi_m \in \mathfrak{g}\,(U_S^\times)$ be representatives of a basis of Čech cohomology classes in $H^1(\Sigma, \mathrm{Ad}(P_g))$, $m = (\mathrm{genus}(\Sigma) - 1)\dim(G)$. Then, for each fixed $w \in U_S$ there exists a unique meromorphic one-form $r(z, w)dz$ on U_S with values in $\mathfrak{g} \otimes \mathfrak{g}$ such that*

(i) *$r(z, w)dz$ is regular on $U_S - \{w\}$ and, as $z \to w$,*

$$r(z, w)dz = \frac{C}{z - w}\,dz + O(1),$$

(ii) *$\mathrm{Ad}(g(z)^{-1})^{(1)} r(z, w)dz$ extends to a holomorphic one-form on $\Sigma - S$.*

(iii) *Let γ be the sum of simple closed curves in U_S encircling counterclockwise the points $p_1 \dots, p_n \in S$ and w. Then*

$$\oint_\gamma \mathrm{tr}^{(1)}\left(\xi_j(z)^{(1)} r(z, w)\right) dz = 0, \qquad j = 1, \dots, m$$

We use here the notation $x^{(j)}$ to denote the action of the linear map x on the j^{th} factor of a tensor product of vector spaces. So, for instance, $\mathrm{tr}^{(1)}(x \otimes y) = \mathrm{tr}(x)\, y$.

This object $r(z, w)dz$, the classical r-matrix, appears in [8] in a description of Poisson brackets and integrals of motions of Hitchin systems.

Lemma 4.2. *Let $r(z,w;g,(\xi_j))$ be the classical r-matrix corresponding to the data $g\in G(U_S^\times)$ and $\xi_1,\dots,\xi_m\in\mathfrak{g}\,(U_S^\times)$. Then, for any $h\in G(U_S)$,*

$$r(z,w;hg,(\mathrm{Ad}(h)\xi_j))=\mathrm{Ad}(h(z))\otimes\mathrm{Ad}(h(w))r(z,w;g,(\xi_j)).$$

For any $n\in G(\Sigma-S)$, $r(z,w;gn,(\xi_j))=r(z,w;g,(\xi_j))$. For any $x_1,\dots,x_m\in\mathfrak{g}\,(U_S)$,

$$r(z,w;g,(\xi_j+x_j))=r(z,w;g,(\xi_j))-\sum_{j=1}^m\omega_j(z)\otimes x_j(w).$$

Finally, for any $y_1,\dots,y_m\in\mathfrak{g}\,(\Sigma-S)$,

$$r(z,w;g,(\xi_j+\mathrm{Ad}(g)y_j))=r(z,w;g,(\xi_j)).$$

Proof: By the uniqueness of r, the proof consists of checking the characterizing properties (i)-(iii), which is straightforward. □

The dependence of $r(z,w)$ on the second argument is given by the following result. Let $\omega_j(z)dz$, $j=1,\dots,m$, be a basis of one-forms in $H^0(\Sigma,K\otimes\mathrm{Ad}(P_g))$ dual (by Serre duality) to the basis (ξ_j).

Thus $\omega_j(z)dz$ is a $\mathfrak{g}$-valued holomorphic one-form on U_S, such that $\mathrm{Ad}(g(z)^{-1})\omega_j(z)dz$ extends to a holomorphic one-form on $\Sigma-S$ and

$$\frac{1}{2\pi i}\oint\mathrm{tr}\,\omega_j(z)\xi_l(z)dz=\delta_{jl}.$$

Lemma 4.3. *For fixed $z\in U_S$, the classical r-matrix $r(z,w)$ is a holomorphic function of $w\in U_S-\{z\}$. Moreover $\mathrm{Ad}(g(w)^{-1})^{(2)}(r(z,w)+\omega_j(z)\otimes\xi_j(w))$ extends to a holomorphic function of w on $\Sigma-(S\cup\{z\})$*

A different characterization of classical r-matrices is as kernels of projections (cf. [6], §1). In the present context, this characterization also holds:

Theorem 4.4. *Let $\mathfrak{g}\,(\Sigma-S,g)$ be the Lie subalgebra of $\mathfrak{g}\,(U_S^\times)$ of functions x, such that $\mathrm{Ad}(g^{-1})x$ extends to $\Sigma-S$. Let $P_+:\mathfrak{g}\,(U_S^\times)\to\mathfrak{g}\,(U_S^\times)$ be the projection onto $\mathfrak{g}\,(U_S)$ in the decomposition*

$$\mathfrak{g}\,(U_S^\times)=\mathfrak{g}\,(U_S)\oplus(\oplus_{j=1}^m\mathbb{C}\,\xi_j)\oplus\mathfrak{g}\,(\Sigma-S,g).\tag{8}$$

Then, for any $x\in\mathfrak{g}\,(U_S^\times)$ and $w\in U_S$,

$$(P_+x)(w)=\frac{1}{2\pi i}\oint_\gamma\mathrm{tr}^{(1)}x(z)^{(1)}r(z,w)dz.$$

Proof: If $x\in\mathfrak{g}\,(U_S)$, then the integrand has, by (i), only a pole at w with residue $\mathrm{tr}^{(1)}(x(w)^{(1)}C)=x(w)$. If x is in the subspace spanned by the ξ_j, the integral is zero by (iii). Finally, if $x(t)=\mathrm{Ad}(g(t))y(t)$ with $y(t)$ holomorphic on $\Sigma-S$, then

$$\oint_\gamma\mathrm{tr}^{(1)}x(t)^{(1)}r(t,w)\,dt=\oint_\gamma\mathrm{tr}^{(1)}y(t)^{(1)}\mathrm{Ad}(g(t)^{-1})^{(1)}r(t,w)\,dt=0,$$

by Stokes' theorem, since γ is the boundary of a region where the integrand is a closed one-form. □

Let us now suppose that we have a local parametrization $\lambda \mapsto [P_{g_\lambda}]$ of the moduli space of stable G-bundles by some region Λ of $\mathbb{C}^m$: for each $\lambda \in \Lambda$, we suppose to have a $g_\lambda \in G(U_S^\times)$ holomorphic in λ, so that, for all $\lambda \in \Lambda$ the m vectors in $\mathfrak{g}\,(U_S)$

$$\xi_{\lambda j} = \frac{\partial g_\lambda}{\partial \lambda_j} g_\lambda^{-1}$$

represent linearly independent classes in $H^1(\Sigma, \mathrm{Ad}(P_{g_\lambda}))$, the tangent space to the moduli space at g_λ.

Let us denote by $r(z, w; \lambda)dz$ the classical r-matrix of Proposition 4.1 with $g = g_\lambda$ and $\xi_j = \xi_{\lambda j}$. Let $\omega_j(z, \lambda)dz$ denote the dual basis in $H^0(\Sigma, K \otimes \mathrm{Ad}(P_g))$. Then we have the "dynamical" classical Yang–Baxter equation

Theorem 4.5. *For all distinct $z_1, z_2, z_3 \in U_S$,*

$$\begin{gathered}[r^{(13)}(z_1, z_3, \lambda), r^{(23)}(z_2, z_3, \lambda)] = \\ [r^{(21)}(z_2, z_1, \lambda), r^{(13)}(z_1, z_3, \lambda)] - [r^{(12)}(z_1, z_2, \lambda), r^{(23)}(z_2, z_3, \lambda)] \\ + \sum \omega_j^{(1)}(z_1, \lambda)\tfrac{\partial}{\partial \lambda_j} r^{(23)}(z_2, z_3, \lambda) - \sum \omega_j^{(2)}(z_2, \lambda)\tfrac{\partial}{\partial \lambda_j} r^{(13)}(z_1, z_3, \lambda).\end{gathered}$$

The notation is that $t^{(ij)} = \sum x^{(i)} y^{(j)}$ if $t = \sum x \otimes y$.

4.2. **An explicit form for the connection.** In this section we give an explicit form of the energy-momentum tensor in terms of local coordinates on the moduli space of G-bundles. It is given in terms of the dynamical classical r-matrix. Thus we consider, as before, a Riemann surface Σ, and an open neighborhood U_S of a finite set $S \subset \Sigma$. Then for each $g \in G(U_S^\times)$ and each choice of a set of vectors $\xi_1, \dots, \xi_m$ in $\mathfrak{g}\,(U_S)^\times$ representing a basis in $H^1(\Sigma, \mathrm{Ad}(P_g))$, we have an r-matrix $r(z, t)$.

The energy-momentum tensor is associated to a point on some open set U_0, on which a local coordinate z is chosen. To simplify the notation, we will simply write z, to denote a point in U_0, which we view as a aubset of $\mathbb{C}$ via the local coordinate. We extend g and ξ_j to maps from $U_S^\times \cup U_0$ by setting $g = 1$ and $\xi_j = 0$ on U_0, and get in this way an r-matrix $r(z, t)$ on $U_S \cup U_0$. By Lemma 4.3, $r(z, t)dz$ for $z, t \in U_0$ is the analytic continuation of

$$\mathrm{Ad}(g(z)^{-1}) \otimes \mathrm{Ad}(g(t)^{-1})(r(z, t) + \sum \omega_j(z) \otimes \xi_j(t))dz$$

from $U_S \times U_S$. We keep the notation $r(z, t)$ for this new r-matrix since it coincides with the old one on $U_S \times U_S$. We will also need $r(z, t)dz$ when $z \in U_0$ and $t \in U_S$. This is the analytic continuation of $\mathrm{Ad}(g(z)^{-1})^{(1)} r(z, t)dz$.

Let $\lambda_1, \dots, \lambda_m$ be local complex coordinates on the moduli space of stable G-bundles. We will fix for $\lambda \in \Lambda \subset \mathbb{C}^m$ an element $g_\lambda \in G(U_S^\times)$ holomorphically depending on λ and representing the corresponding isomorphism class of G-bundles. The image of the coordinate vector fields $\partial/\partial\lambda_j$ are the classes in the tangent space $H^1(\Sigma, \mathrm{Ad}(P_{g_\lambda}))$ at $[P_{g_\lambda}]$ represented by the vectors in $\mathfrak{g}\,(U_S^\times)$

$$\xi_{\lambda,j}(t) = \partial_{\lambda_j} g_\lambda(t)\, g_\lambda(t)^{-1}.$$

We also choose a holomorphic lift $\hat{g}_\lambda = [(g_\lambda^{\mathrm{reg}}, z_\lambda)] \in G(U_S^\times)$, which amounts to choosing a trivialization of the vector bundle of which conformal blocks are sections. The choice of this lift is encoded in functions $a_j(\lambda)$ defined by

$$\begin{aligned} D_{\xi_j} f(\hat{g}_\lambda) &= -(\partial_{\lambda_j} + k a_j(\lambda)) f(\hat{g}_\lambda) \\ &=_{\mathrm{def}} -\nabla_{\lambda_j} f(\hat{g}_\lambda), \end{aligned}$$

for any local holomorphic function f. Explicitly,

$$a_j(\lambda) = \frac{1}{2\pi i}\int_{U_S} \mathrm{tr}(\partial(\partial_{\lambda_j} g_\lambda^{\mathrm{reg}} g_\lambda^{\mathrm{reg}-1})\bar\partial g_\lambda^{\mathrm{reg}} g_\lambda^{\mathrm{reg}-1}) - \partial_{\lambda_j}\log(z_\lambda).$$

Definition: Let $r(z,t,\lambda)$ be the r-matrix corresponding to the data g_λ, $\xi_{\lambda,j}$. Let

$$r(z,t,\lambda) = \frac{C}{z-t} + r_0(z,\lambda) + O(t-z) + (t-z)r_1(z,\lambda) + O((t-z)^2),$$

be the Laurent expansion of the classical r-matrix at its pole.

Let $q_1(z,\lambda) = [r_0(z,\lambda)]$, where $[\quad]: \mathfrak{g}\otimes\mathfrak{g} \to \mathfrak{g}$ is the Lie bracket on $\mathfrak{g}$.

Let $q_2(z,\lambda) = \mu(r_1(z,\lambda))$, where $\mu : \mathfrak{g}\otimes\mathfrak{g} \to \mathbb{C}$ is the invariant symmetric bilinear form $x\otimes y \mapsto \mathrm{tr}(xy)$.

One essential ingredient is the ℓ-operator. Its semiclassical counterpart is a higher genus version of the Lax operator in the r-matrix formulation of classical integrable systems.

Definition: Let, for $z \in U_0$,

$$q_3(z,\lambda) = \frac{-1}{2\pi i}\oint_\gamma \mathrm{tr}^{(2)}\big((r(z,t,\lambda) + \omega_j(z,\lambda)\otimes\xi_{\lambda j}(t))(dg_\lambda(t)g_\lambda(t)^{-1})^{(2)}\big). \tag{9}$$

The ℓ-*operator* is the differential operator in λ with values in $\mathfrak{g}\otimes\mathrm{End}(V)$:

$$\hat\ell(z) = \sum_{j=1}^m \omega_j(z,\lambda)^{(0)}\nabla_{\lambda_j} + kq_3(z,\lambda)^{(0)} + \sum_{j=1}^n r(z,p_j,\lambda)^{(0j)},$$

where the factors in $\mathfrak{g}\otimes\mathrm{End}(V) = \mathfrak{g}\otimes\mathrm{End}(V_1)\otimes\cdots\otimes\mathrm{End}(V_n)$ are numbered from 0 to n. The "spectral parameter" z runs over U_0.

Proposition 4.6. *Let $z\in U_0$ and $\lambda\in\Lambda$. Set $\kappa = k + h^\vee$. Then*

(i) $J_x(z)u(\hat g_\lambda) = -\mathrm{tr}^{(0)}(x^{(0)}\hat\ell(z))\, u(\hat g_\lambda)$

(ii) $T(z)u(\hat g_\lambda) = \frac{1}{2\kappa}\big(\mathrm{tr}^{(0)}\hat\ell(z)^2 + \mathrm{tr}^{(0)}q_1(z,\lambda)^{(0)}\hat\ell(z) + kq_2(z,\lambda)\big)u(\hat g_\lambda)$

Proof: We introduce the notation $\hat\ell_x(z) = \mathrm{tr}^{(0)}x^{(0)}\hat\ell(z)$.

(i) By definition, $J_x(z)\,u(\lambda) = D_{x_{-1}(z)}i_z^{-1}u(i_z\hat g_\lambda)$, where $x_{-1}(z)$ is the map $t\mapsto x/(t-z)$ for $t\in U_0$ and is equal to zero on U_S.

To compute this, we have to decompose $x_{-1}(z)$ according to the decomposition (8): $x_{-1}(z) = y_+ + y_0 + y_-$.

The first component is $y_+ := P_+x_{-1}(z)$. The contribution of this piece to $J_x(z)u$ is $-\sum_{j=1}^n y_+(p_j)^{(j)}u(\hat g_\lambda)$.

By Theorem 4.4, $y_+(p_j)$ is given by a contour integral over γ. Since x_+ vanishes except on U_0, only the component $\gamma_z = \gamma\cap U_0$ contributes to the integral:

$$y_+(p_j) = \frac{1}{2\pi i}\oint_{\gamma_z}\mathrm{tr}^{(1)}\frac{x^{(1)}}{w-z}r(w,p_j,\lambda)dw = \mathrm{tr}^{(1)}x\,r(z,p_j).$$

The second component is

$$y_0(t)=\sum_{j=1}^{m}\frac{1}{2\pi i}\oint_{\gamma_z}\mathrm{tr}\left(\frac{x}{w-z}\,\omega_j(w)\,dw\right)\,\xi_{\lambda j}(t)=\sum_{j=1}^{m}\mathrm{tr}(x\,\omega_j(z))\xi_{\lambda j}(t).$$

Since $D_{\xi_{\lambda_j}} i_z^{*-1}u\circ i_z=-\nabla_{\lambda_j}u$, we get

$$D_{y_0}i_z^{*-1}u(i_z\hat g_\lambda)=-\sum \mathrm{tr}(x\,\omega_j(z,\lambda))\partial_{\lambda_j}u(\hat g_\lambda).$$

We turn to the third component. By Lemma 2.4,

$$D_{y_-}i_z^{*-1}u(i_z\hat g_\lambda)=-\frac{k}{2\pi i}\oint_{\gamma\cup\gamma_z}\mathrm{tr}(di_z(g_\lambda)i_z(g_\lambda)^{-1}y_-)u(\hat g_\lambda). \tag{10}$$

Since $i_z(g_\lambda)$ is trivial on U_0, the integral reduces to an integral over γ. By Theorem 4.4, for $t\in U_S$,

$$\begin{aligned} y_-(t) &= -y_0(t)-y_+(t)\\ &= -\sum_j \mathrm{tr}(\omega_j(z,\lambda)\,x)\xi_{\lambda j}(t)-\mathrm{tr}^{(1)}x^{(1)}r(z,t,\lambda). \end{aligned}$$

Therefore,

$$\begin{aligned} D_{y_-}i_z^{*-1}u(i_z\hat g_\lambda) &= \frac{1}{2\pi i}\oint_\gamma \mathrm{tr}\otimes\mathrm{tr}\big(x\otimes dg_\lambda(t)g_\lambda(t)^{-1}\\ &\quad (r(z,t,\lambda)+\omega_j(z,\lambda)\otimes\xi_{\lambda j}(t))\big)u(\hat g_\lambda)\\ &\quad =-k\,\mathrm{tr}(x\,q_3(z,\lambda))u(\hat g_\lambda). \end{aligned}$$

These terms taken together give $-\hat\ell$ after properly renumbering the factors, as claimed.
(ii) Let $x\in\mathfrak g$, and $y=x_{-1}(z)$ (we will later take $x=b_j$ and sum over j). Thus

$$y(t)=\begin{cases}\frac{x}{t-z}, & t\in U_0\\ 0, & t\in U_S.\end{cases}$$

Let us decompose $y=y_++y_0+y_-$. Then

$$D_y^2 i_z^{*-1}u=D_{y_+}D_y i_z^{*-1}u+D_{y_0}D_y i_z^{*-1}u+D_{y_-}D_y i_z^{*-1}u.$$

Let us consider the three terms one after the other. Let us introduce the notation $\hat g_{\lambda,s}=\exp(-sy_+)i_z(\hat g_\lambda)$. We then have, by definition,

$$\begin{aligned} D_{y_+}D_y i_z^{*-1}u(i_z\hat g_\lambda) &= \left.\frac{d}{ds}\right|_{s=0} D_y i_z^{*-1}u(\hat g_{\lambda,s})\\ &= \left.\frac{d}{ds}\right|_{s=0}\big(-\hat\ell_x(z;\hat g_{\lambda,s},(\partial_{\lambda_j}g_{\lambda,s}g_{\lambda,s}^{-1}))i_z^{*-1}u(\hat g_{\lambda,s}) \qquad (11)\\ &\quad -\frac{k}{2\pi i}\oint_{\gamma_z}\mathrm{tr}(dg_{\lambda_s}g_{\lambda,s}^{-1}y_-)u(\hat g_\lambda)\big)\\ &= A+B, \end{aligned}$$

where we have written the dependence of $\hat\ell_x(z)$ on the data explicitly. The last term B is the contribution of γ_z to the integral (10), which now does not vanish, since, on U_0,

$g_{\lambda,s} = \exp(-sy_+) \neq 1$. We can give a more explicit formula for this term:

$$\begin{aligned} B &= -\frac{d}{ds}\bigg|_{s=0} \frac{k}{2\pi i} \oint_{\gamma_z} \mathrm{tr}(dg_{\lambda_s} g_{\lambda,s}^{-1} y_-) \\ &= \frac{k}{2\pi i} \oint_{\gamma_z} \mathrm{tr}(dy_+ y_-) \\ &= k\,\mathrm{tr}(y_+'(z)x), \end{aligned}$$

by the residue theorem, since $y_-(t) = x/(t-z) + O(1)$, when $t \to z$. Now,

$$\begin{aligned} y_+(t) &= \frac{x}{t-z} + \mathrm{tr}^{(1)} x^{(1)} r(z,t,\lambda) \\ &= \mathrm{tr}^{(1)} x^{(1)} r_0(z,\lambda) + (t-z)\mathrm{tr}^{(1)}(x^{(1)} r_1(z,\lambda)) + O((t-z)^2). \end{aligned}$$

It follows that $y_+'(z) = \mathrm{tr}^{(1)}(x^{(1)} r_1(z,\lambda))$. Thus

$$k\,\mathrm{tr}(y_+'(z)x) = k\,\mathrm{tr}\otimes\mathrm{tr}(x\otimes x\, r_1(z,\lambda)).$$

If $x = b_j$, the sum over j of this expression is $k\mu(r_1(z,\lambda))$.

To deal with the first term in (11), we need to know how $\hat\ell$ depends on s.

Lemma 4.7. *Let $g_\lambda \in \hat G(U_R^\times)$ for all $\lambda \in \Lambda$ and $h \in G(U_R)$. Let $\tilde g_\lambda = hg_\lambda$, $\xi_{\lambda j} = \partial_{\lambda_j} g_\lambda g_\lambda^{-1}$ and $\tilde\xi_{\lambda j} = \partial_{\lambda_j}\tilde g_\lambda \tilde g_\lambda^{-1}$. Then*

$$\hat\ell_x(z;\tilde g_\lambda,(\tilde\xi_{\lambda j})) = \prod_{l=1}^n h(p_l)^{(l)} \hat\ell_{\mathrm{Ad}(h(z)^{-1})x}(z;g_\lambda,(\xi_{\lambda j})) \prod_{l=1}^n h(p_l)^{-1(l)}.$$

This follows from Lemma 4.2, the following transformation rule for the holomorphic one-forms $\omega_j(z;g_\lambda,(\xi_{\lambda j}))$:

$$\omega_j(z;\tilde g_\lambda,(\tilde\xi_{\lambda j})) = \mathrm{Ad}(h(z))\omega_j(z;g_\lambda,(\xi_{\lambda j})),$$

and the fact that if $\hat g_\lambda \to h\hat g_\lambda$,

$$\begin{aligned} a_j(\lambda) &\to a_j(\lambda) + \frac{1}{2\pi i}\oint_\gamma \mathrm{tr}\big(\xi_{\lambda j}(t) h^{-1} dh(t)\big), \\ q_3(z,\lambda) &\to q_3(a,\lambda) + \sum_j \omega_j(z,\lambda)\frac{1}{2\pi i}\oint_\gamma \mathrm{tr}\big(\xi_{\lambda j}(t) h^{-1} dh(t)\big), \end{aligned}$$

leading to a cancellation.

In our case, $R = S\cup\{z\}$ and $h_= \exp(-sy_+)$. Using the invariance $u(hg) = h\,u(g)$, we get

$$\begin{aligned} A &= -\frac{d}{ds}\bigg|_{s=0} \prod_{j=1}^n \exp(-sy_+(p_j))^{(j)} \hat\ell_{\mathrm{Ad}(\exp\, sy_+(z))x}(z) i_z^{*-1} u(\hat g_\lambda) \\ &= (-\hat\ell_{[y_+(z),x]}(z) + \sum_{l=1}^n y_+(p_l)^{(l)}) u(\hat g_\lambda) \\ &= -\mathrm{tr}^{(0)}\big([\mathrm{tr}^{(-1)} x^{(-1)} r_0(z,\lambda)^{(-1,0)}, x^{(0)}]\hat\ell(z)\big) u(g_\lambda) \\ &\quad + \sum_j \mathrm{tr}^{(0)}\big(x^{(0)} r(z,t,\lambda)^{(0j)}\big)\hat\ell_x(z)\, u(g_\lambda). \end{aligned}$$

We have used here the fact that, for $s \in U_0$,

$$\begin{aligned} y_+(s) &= \frac{1}{2\pi i}\oint_{\gamma_z} \mathrm{tr}^{(1)} \frac{x^{(1)}}{t-z} r(t,s)\,dt \\ &= \frac{x}{s-z} + \mathrm{tr}^{(1)} x^{(1)} r(z,s). \end{aligned}$$

In particular $y_+(z) = \mathrm{tr}^{(1)} x^{(1)} r_0(z,\lambda)$. To avoid renumbering the factors, we gave the number -1 to the first factor or r_0. Note that, taking $x = b_j$, we have

$$\sum_j [\mathrm{tr}^{(-1)} b_j^{(-1)} r_0(z,\lambda)^{(-1,0)}, b_j^{(0)}] = -[r(z,\lambda)]^{(0)}.$$

This follows from the identity $\mathrm{tr}(b_j a)[b, b_j] = [b, a]$, valid for any $a \otimes b$.

The second term in the decomposition is treated easily:

$$D_{y_0} D_y i_z^{*-1} u(i_z \hat{g}_\lambda) = \sum_{j=1}^{m} \mathrm{tr}(x\,\omega_j(z,\lambda)) \nabla_{\lambda_j}(\hat{\ell}_x(z) u(g_\lambda))$$

We turn to the third term. Since $D_y i_z^{*-1} u$ is invariant under the right action of $G(\Sigma - (S \cup \{z\}))$, we have, by Lemma 2.4,

$$D_{y_-} D_y i_z^{*-1} u(i_z \hat{g}_\lambda) = -\frac{k}{2\pi i}\oint_\gamma \mathrm{tr}(dg_\lambda(t) g_\lambda(t)^{-1} y_-(t)) D_y i_z^{*-1} u(i_z \hat{g}_\lambda).$$

The integral is actually over $\gamma \cup \gamma_z$, but the integral over γ_z vanishes, since g_λ is extended by 1 on $U_0^\times$. Since, for $t \in U_S$,

$$y_-(t) = -\mathrm{tr}^{(1)} x^{(1)} r(z,t,\lambda) - \sum_j \mathrm{tr}(x\,\omega_j(z,\lambda)) \xi_{\lambda j}(t),$$

it follows that

$$D_{y_-} D_y i_z^{*-1} u(i_z \hat{g}_\lambda) = \mathrm{tr}(x\, q_3(z,\lambda)) \hat{\ell}_x(z,\lambda) u(\hat{g}_\lambda).$$

We then take all terms together, set $x = b_j$ and sum over j. □

4.3. **Transformation properties.** In this section we compute the dependence of the various objects we introduced on the choices of coordinates and trivializations. This computations will show that the connection is indeed well defined.

Lemma 4.8. *The dependence of q_1, q_2, q_3 on the choice of coordinate on U_0 is*

$$\begin{aligned} q_1^w(w,\lambda) &= q_1^z(z,\lambda)\frac{dz}{dw} \\ q_2^w(w,\lambda) &= q_2^z(z,\lambda)\left(\frac{dz}{dw}\right)^2 + \frac{\dim(G)}{6}\{z,w\} \\ q_3^w(w,\lambda) &= q_3^z(z,\lambda)\frac{dz}{dw} \end{aligned}$$

To prove this lemma one uses that $r(z,t)$ depends on coordinates locally as a one-form in the first argument and as a function in the second. Then one compares the Laurent expansions of r in two different coordinates. This gives the formulae for q_1, q_2. As for q_3, the statement is obvious.

Let us from now on assume that all coordinates we consider are part of a flat atlas, so that Schwarzian derivatives do not appear. Then q_1dz, q_3dz have an analytic continuation to holomorphic 1-forms away from S, and q_2dz^2 has an analytic continuation to a holomorphic quadratic differential away form S. The singularity at S can be arbitrary. However, using the transformation behavior of r, we deduce

Proposition 4.9. *q_1dz, q_2dz^2, q_3dz extend to holomorphic (quadratic) differentials on $\Sigma - S$. Moreover, for $z \in U_S^\times$, and if we denote the derivative with respect to z by a prime,*

$$\mathrm{Ad}(g_\lambda(z))q_1(z,\lambda)dz + 2h^\vee dg_\lambda(z)g_\lambda(z)^{-1} - \sum[\omega_j(z,\lambda), \xi_{\lambda j}(z)]dz,$$
$$\big(q_2(z,\lambda) - \sum \mathrm{tr}(\omega_j(z,\lambda)\xi_j'(z,\lambda)) + \mathrm{tr} g_\lambda(z)^{-1}g_\lambda'(z)q_1(z,\lambda)$$
$$-h^\vee \mathrm{tr}\big((g_\lambda'(z)g_\lambda(z)^{-1})^2)\big)dz^2,$$
$$\mathrm{Ad}(g_\lambda(z))q_3(z,\lambda)dz + dg_\lambda(z)g_\lambda(z)^{-1},$$

extend to holomorphic (quadratic) differentials q_1^Sdz, $q_2^Sdz^2$, q_3^Sdz on U_S.

Remark. $q_1^S = [r_0]$ and $q_2^S = \mu(r_1)$ where r_0 and r_1 are the Laurent coefficients of the r-matrix on U_S. Also, q_3^S is given by the integral formula (9) but with an integration contour that encircles z, as well as the points p_j.

Finally, we need the transformation behavior of $\hat\ell$ which can be readily deduced from the one of r and q_3.

Proposition 4.10. *$\hat\ell(z)dz$ has an analytic continuation as a function of z to a one-form on $\Sigma - S$. Moreover, for $z \in U_S^\times$,*

$$\hat\ell_S(z) = \mathrm{Ad}(g_\lambda(z))^{(0)}\hat\ell(z)dz + kdg_\lambda(z)g_\lambda(z)^{-1(0)}$$

extends to a meromorphic one-form on U_S with at most simple poles on S.

These results prove in particular part (i) of Proposition 3.2. Part (ii) follows by computing the left derivative (or the right derivative) along the lines of the proof of Proposition 4.6 (i), i.e., by decomposing y into the three components and using the invariance properties of conformal block to express the action in terms of ℓ-operators.

Let us put everything together. We have the second order differential operator $A(z)$ in the variables λ_j acting on V-valued functions, and defined a priori for $z \in U_0$:

$$A(z) = \frac{1}{2\kappa}\big(\mathrm{tr}^{(0)}\hat\ell(z)^2 + \mathrm{tr}^{(0)}q_1(z,\lambda)^{(0)}\hat\ell(z) + kq_2(z,\lambda)\big)$$

Corollary 4.11. *$A(z)dz^2$ extends (given a flat structure) to a quadratic differential on $\Sigma - S$. Moreover, if $z \in U_S^\times$,*

$$A_S(z) = A(z) + \mathrm{tr}^{(0)}((g_\lambda'(z)g_\lambda(z)^{-1})^{(0)}\hat\ell(z)) + \frac{k}{2}\mathrm{tr}((g_\lambda'(z)g_\lambda(z)^{-1})^2)$$

extends to a meromorphic quadratic differential on U_S with at most poles of second order on S.

In particular, this proves Proposition 3.7.

5. Moving points

We describe here the KZB equations that correspond to moving the points and the tangent vectors, but keeping the complex structure on Σ fixed.

5.1. **Fixing the complex structure.** Let us thus fix a Riemann surface (Σ, I) of genus h with complex structure I and let B_n be the moduli space of data $(q_i, w_i)_{i=1}^n$ consisting of n distinct points q_i on Σ and n non-zero tangent vectors $w_i \in T_{q_i}\Sigma$, modulo the natural action of the group of conformal automorphisms of Σ. If $n \geq 1$ ($n \geq 2$ for genus zero), B_n is the smooth algebraic variety

$$B_n = T^\times(\Sigma^n - \cup_{i<j}\{x_i = x_j\})/\mathrm{Aut}(\Sigma),$$

where $T^\times$ denotes the complement in the holomorphic tangent bundle of the set of vectors with at least one vanishing component.

We have an embbedding $j : B_n \hookrightarrow \hat{M}_{h,n}$ sending the class of (q_i, w_i) to the class of $\phi_*(I)$ for any diffeomorphism $\phi : \Sigma \to \Sigma$ such that $\phi(q_i) = p_i$ and $\phi_*(q_i)w_i = v_i$ for all $i = 1, \dots, n$.

Therefore the connection on conformal blocks induces a connection on the pull-back of the projectivized vector bundle of conformal blocks on B_n. To describe this connection, we need to study the tangent map j_* of j. Let $(\dot{q}_i, \dot{w}_i)$ be a tangent vector at (q_i, w_i) and let $(q_i(s), w_i(s))$, $|s| < \epsilon$, represent a curve in B_n with tangent vector $(\dot{q}_i, \dot{w}_i)$ at $s = 0$. Suppose ϕ is a diffeomorphism sending (q_i, w_i) to (p_i, v_i), and let $z : U_S \to \sqcup_{j=1}^n \mathbb{C}$ be a local complex coordinate for $\phi_*(I)$ on some neighborhood U_S of S as above. This coordinate pulls back to a local flat coordinate $z \circ \phi$ on a neighborhood of $\{q_1, \dots, q_n\}$. Let $z_i(s) = z \circ \phi(q_i(s))$, $\eta_i(s)\frac{d}{dz} = z \circ \phi_*(q_i(s))w_i$ be the expression of our family with respect to this local coordinate. By construction, $z_i(0) = 0$, $\eta_i(0) = 1$. For $|s| < \epsilon$, let us choose a diffeomorphism ϕ_s of Σ such that $z(\phi_s(q)) = \eta_i(s)^{-1}(z(\phi(q)) - z_i(s))$ for q close to q_i, and such that $\phi_0 = \phi$. Such diffeomorphisms can be easily constructed by taking $\phi_s = \phi$ except on some neighborhood of the points q_i and choosing a suitable interpolation in an annular region around the points. The diffeomorphism ϕ_s sends $(q_i(s), w_i(s))$ to (p_i, v_i), so it can be used to define $j(q_i(s), w_i(s)) = \mathrm{cls}(\phi_{s*}(I))$. Then z is still a local complex coordinate in a sufficiently small neighborhood of S for all complex structures $\phi_{s*}(I))$, and we have a curve $(\phi_{s*}(I), z)$ in the enlarged moduli space (see 3.3).

The tangent vector $\dot{I}$ to the curve $\phi_{s*}(I)$ corresponds to a Beltrami differential $\mu = -\bar{\partial}\xi$ where ξ is the vector field on Σ such that $\xi(\phi(x)) = \frac{d}{ds}|_{s=0}\phi_s(x)$. This vector field is holomorphic on a neighborhood U_S of S, thus $\mu = 0$ on U_S. The tangent vector $(\dot{I}, \dot{z} = 0)$ to the curve in the enlarged moduli space is then represented by the holomorphic vector field $\zeta = \xi$ on $U_S^\times$. Indeed, we can write $\mu = \bar{\partial}\zeta_S$ on U_S and $\mu = \bar{\partial}\zeta_\infty$ on the complement of S with $\zeta_S = 0$ and $\zeta_\infty = -\xi$. Thus $\zeta = \zeta_S - \zeta_\infty = \xi$.

Let us summarize.

Proposition 5.1. *Let us fix some coordinate $z : U_S \to \sqcup_{j=1}^n \mathbb{C}$ on some neighborhood of S and let $z_1, \dots, z_n, \eta_1, \dots, \eta_n$ be the coordinates of a point in B_n in neighborhood of (p_i, v_i). Then the covariant derivative in the direction of a tangent vector $(\dot{z}_i, \dot{\eta}_i)$ is given by*

$$\begin{aligned}\nabla_{\dot{z}_i,\dot{\eta}_i} u(z_i, \eta_i, \hat{g}_\lambda) &= \sum \dot{z}_i \frac{\partial}{\partial z_i} u(z_i, \eta_i, \hat{g}_\lambda) \\ &\quad + \sum \dot{\eta}_i \frac{\partial}{\partial \eta_i} u(z_i, \eta_i, \hat{g}_\lambda) \\ &\quad + \frac{1}{2\pi i} \oint_\gamma T^S(z)\zeta(z) dz\, u(z_i, \eta_i, \hat{g}_\lambda),\end{aligned}$$

where $\zeta = -(\dot\eta_i z + \dot z_i)\frac{d}{dz}$

The fact that T^S rather than T appears in this formula is due to the choice of describing the bundle by a function g_λ which is a fixed function of the coordinate z when we move the points, whereas in (4) g_λ is fixed in the coordinate vanishing at S. The difference is given in terms of J and cancels the terms in Proposition 3.7. The details are left to the reader.

We now give a more explicit formula for the connection in this case. We keep the notation of the previous proposition, and express our connection in terms of the r-matrix $r(z,w,\lambda)$, $z,w\in U_S$, its constant term r_0, defined by

$$r(z,w,\lambda)=\frac{C}{z-w}+r_0(z,\lambda)+O(w-z),$$

and the one-form $q(z,\lambda)=\sum\omega_\nu(z,\lambda)a_\nu(\lambda)+q_3^S(z,\lambda)$, where $a_\nu(\lambda)$ depend on the choice of local trivialization of the $\mathbb{C}^\times$-bundle $\hat G$, see 4.2, and q_3^S may be characterized by the properties:

(i) $q_3^S(z,\lambda)$ is holomorphic on U_S.
(ii) $\mathrm{Ad}(g_\lambda(z)^{-1})q_3^S(z,\lambda)dz-g_\lambda(z)^{-1}dg_\lambda(z)$ extends to a holomorphic one-form on $\Sigma-S$.
(iii) $\oint_\gamma \mathrm{tr}(q_3^S\xi_\nu)dz=0$ for all $\nu=1,\dots,m$.

Note that q_3^S can also be expressed in terms of the r-matrix, and g_λ, see the remark after Proposition 4.9.

We will denote by $\mathrm{Cas}(V_j)$ the value of the central Casimir element $\sum b_i^2$ on the irreducible representation V_i. Evaluating the contour integrals, we get

Theorem 5.2. *The connection on the space of conformal blocks restricted to B_n is given by the formula:*

$$\nabla=\sum dz_i\nabla_{z_i}+\sum d\eta_i\nabla_{\eta_i}$$

where

$$\begin{aligned}\nabla_{z_i} &= \frac{\partial}{\partial z_i}-\frac{1}{\kappa}\Bigg(\sum_{\nu=1}^{n}\omega_\nu(z_i,\lambda)^{(i)}\frac{\partial}{\partial\lambda_\nu}+kq(z_i,\lambda)^{(i)}\\ &\qquad +r_0(z_i,\lambda)^{(ii)}+\sum_{j:j\neq i}r(z_i,z_j,\lambda)^{(ij)}\Bigg),\end{aligned}$$

and

$$\nabla_{\eta_i}=\frac{\partial}{\partial\eta_i}-\frac{\mathrm{Cas}(V_i)}{2\kappa}$$

The next result is that this connection is flat. In fact, more strongly, we have the following result for the differential operators $\nabla_{z_i},\nabla_{\eta_i}$ acting on any functions of z,η,λ, not just conformal blocks.

Theorem 5.3. *The differential operators $\nabla_{z_i},\nabla_{\eta_i}$ with coefficients in $U(\mathfrak{g})^{\otimes n}$ defined by the formulae in Theorem 5.2 commute with each other for any complex value of $k\neq -h^\vee$.*

In particular, we may take the limit $k\to -h^\vee$ and obtain, for each (z_i,η_i), a set of commuting differential operators in the coordinates λ, and forming a generalization of the integrable Gaudin model:

Corollary 5.4. *For each $z_1, \dots, z_n$, the differential operators*

$$H_i = \sum_{\nu=1}^{n} \omega_\nu(z_i, \lambda)^{(i)} \frac{\partial}{\partial \lambda_\nu} - h^\vee q(z_i, \lambda)^{(i)} + r_0(z_i, \lambda)^{(ii)} + \sum_{j:j\neq i} r(z_i, z_j, \lambda)^{(ij)},$$

commute with each other.

5.2. **Proof of Theorem 5.2.** The only thing that we have to prove is that $\nabla_{z_j}\nabla_{z_l} = \nabla_{z_l}\nabla_{z_j}$. We first remark that this follows from the universal (i.e., independent of the representations V_i) identities:

$$\frac{\partial A_1}{\partial z_2} - \frac{\partial A_2}{\partial z_1} = 0, \tag{12}$$

$$\kappa\left(\frac{\partial r^{(12)}}{\partial z_2} - \frac{\partial r^{(21)}}{\partial z_1}\right) + [A_1 + r^{(12)}, A_2 + r^{(21)}] = 0, \tag{13}$$

$$\begin{aligned} &[A_1, r^{(23)}] + [r^{(13)}, A_2] + [r^{(13)}, r^{(23)}] \\ &+[r^{(12)}, r^{(23)}] - [r^{(21)}, r^{(13)}] = 0. \end{aligned} \tag{14}$$

Here we have used the abbreviations $r^{(ij)} = r(z_i, z_j, \lambda)^{(ij)} \in \mathfrak{g} \otimes \mathfrak{g}$ and

$$A_j = \sum \omega_\nu(z_j, \lambda)^{(j)} \frac{\partial}{\partial \lambda_\nu} + k\, q(z_j, \lambda)^{(j)} + r_0(z_j, \lambda)^{(jj)}.$$

The differential operators A_j have coefficients in the universal enveloping algebra of $\mathfrak{g}$.

In the last of these identities we recognize the dynamical classical Yang–Baxter equation, Theorem 4.5. The first identity follows trivially from the fact that A_i is independent of z_j, if $i \neq j$. What is left to prove is the identity (13). The left hand side of this identity consists of three parts that, as we now show, vanish separately: the first part is the homogeneous first order part of this first order differential operator. The second part is of zero order as a differential operator and is proportional to k. The third part is of zero order and independent of k.

The vanishing of these three parts is the respective contents of the next three lemmata. We write the formulae in an abbreviated notation: thus $r^{(ij)}$, $q^{(i)}$, $r_0^{(ii)}$, $\omega_\nu^{(i)}$ stand for $r^{(ij)}(z_i, z_j, \lambda)$, $q^{(i)}(z_i, \lambda)$, and so on.

Lemma 5.5. *For all $\mu = 1, \dots, m$,*

$$\sum_\nu \omega_\nu^{(1)} \frac{\partial \omega_\mu^{(2)}}{\partial \lambda_\nu} - \sum_\nu \frac{\partial \omega_\mu^{(1)}}{\partial \lambda_\nu} \omega_\nu^{(2)} + [\omega_\mu^{(1)}, r^{(21)}] + [r^{(12)}, \omega_\mu^{(2)}] = 0.$$

Proof: The left-hand side is regular at $z_1 = z_2$: the poles cancel by the invariance property $[C, x^{(1)} + x^{(2)}] = 0$ valid for any $x \in \mathfrak{g}$. Let us consider this left-hand side as a function of z_1. We use the properties of the r-matrices: we know that $\mathrm{Ad}(g(z_1)^{-1})^{(1)} r^{(12)} dz_1$ extends to a holomorphic one-form on $\Sigma - S$ and that $\mathrm{Ad}(g(z_1)^{-1})^{(1)}(r^{(21)} + \sum \xi_\nu^{(1)} \omega_\nu^{(2)})$ extends to a holomorphic function on $\Sigma - S$. Also, $\mathrm{Ad}(g(z_1)^{-1})^{(1)} \omega^{(1)} dz_1$ extends to a $\mathfrak{g}$-valued holomorphic one-form on $\Sigma - S$. By putting these things together, we see that acting with $\mathrm{Ad}(g(z_1)^{-1})^{(1)}$ on the left-hand side of the identity yields a holomorphic one-form in z_1 on $\Sigma - S$. A similar conclusion holds for z_2 and the second factor. Hence the left-hand side has the form

$$\sum_{\mu,\nu} B_{\mu,\nu} \omega_\mu^{(1)} \omega_\nu^{(2)}. \tag{15}$$

The constant coefficients $B_{\mu,\nu}$ can be calculated as contour integrals over $\gamma\times\gamma$ of the trace of the left-hand side times $\xi_\mu^{(1)}\xi_\nu^{(2)}$. The latter integral can be evaluated using the defining property (iii) of the r-matrix and the definition of the one-forms ω_ν. The result is that the integral vanishes for all μ,ν. Hence (15) is zero and the claim follows. □

Lemma 5.6.

$$\frac{\partial r^{(12)}}{\partial z_2}-\frac{\partial r^{(21)}}{\partial z_1}+\sum_\nu\left(\omega_\nu^{(1)}\frac{\partial q^{(2)}}{\partial\lambda_\nu}-\frac{\partial q^{(1)}}{\partial\lambda_\nu}\omega_\nu^{(2)}\right)+[q^{(1)},r^{(21)}]+[r^{(12)},q^{(2)}]=0.$$

Proof: This lemma is proved in a similar way as the previous one: one uses the transformation and analytic continuation properties of the various objects, to show that the left-hand side is of the form (15) and then shows that the coefficients $B_{\mu,\nu}$ vanish, by computing them as contour integrals. The calculation reduces to

$$B_{\mu,\nu}=\frac{\partial a_\mu}{\partial\lambda_\nu}-\frac{\partial a_\nu}{\partial\lambda_\mu}-\frac{1}{2\pi i}\oint_\gamma\mathrm{tr}(d\xi_\mu\xi_\nu).$$

But this is zero by Proposition 2.3 and the definition of a_ν in terms of D_{ξ_ν}. □

Lemma 5.7.

$$\begin{aligned}h^\vee\left(\frac{\partial r^{(12)}}{\partial z_2}-\frac{\partial r^{(21)}}{\partial z_1}\right)+[r_0^{(11)},r^{(21)}]+[r^{(12)},r_0^{(22)}]+[r^{(12)},r^{(21)}]\\+\textstyle\sum_\nu\left(\omega_\nu^{(1)}\frac{\partial}{\partial\lambda_\nu}(r_0^{(22)}+r^{(21)})-\omega_\nu^{(2)}\frac{\partial}{\partial\lambda_\nu}(r_0^{(11)}+r^{(12)})\right)=0.\end{aligned}$$

Proof: This is a degenerate case of the dynamical classical Yang–Baxter equation. Let $m_{13},\, m_{23}\in\mathrm{Hom}(U\mathfrak{g}^{\otimes 3},U\mathfrak{g}^{\otimes 2})$ be the linear map on the third tensor power of the universal enveloping algebra such that

$$m_{13}(x\otimes y\otimes z)=xz\otimes y,\qquad m_{23}(x\otimes y\otimes z)=x\otimes yz.$$

If $t(z_1,z_2,z_3)\in\mathfrak{g}^{\otimes 3}\subset U\mathfrak{g}^{\otimes 3}$ denotes the difference between the two sides of the dynamical classical Yang–Baxter equation, then the identity claimed in the lemma is equivalent to

$$m_{13}t(z_1,z_2,z_1)+m_{23}t(z_1,z_2,z_2)=0,$$

and thus follows from Theorem 4.5. The straightforward details are left to the reader. Here we will only explain the appearance of the dual Coxeter number in the formula. One of the terms that appear by taking the limit $z_3\to z_1$ in the Yang–Baxter tensor is

$$-m_{13}[C^{(13)},\partial_{z_3}r^{(23)}]=-\sum b_i[b_i,y_j]\otimes x_j,\qquad \text{if } \partial_{z_3}r(z_2,z_3,\lambda)=\sum x_j\otimes y_j.$$

The dual Coxeter number is half the value of the Casimir element in the adjoint representation. It appears in this formula since, for any $x\in\mathfrak{g}$, we have the identity in the universal enveloping algebra

$$\begin{aligned}\sum b_i[b_i,x]&=\tfrac12\sum\left([b_i^2,x]+[b_i,[b_i,x]]\right)\\&=\tfrac12(0+2\,h^\vee x)\\&=h^\vee x.\end{aligned}$$

□

6. Concluding remarks

In this paper we have written the Knizhnik–Zamolodchikov–Bernard equations in arbitrary local coordinates, and checked the independence of choices.

We have addressed the question of flatness of the connection only in the situation where we move the marked points by keeping the complex structure fixed. In fact this connection is supposed to be flat (on the projectivization of the vector bundle of conformal blocks) on the whole moduli space. In the case of no marked points, Hitchin's proof [24] of the flatness goes as follows. The principal symbol of the covariant derivatives ∇_j in the direction of some coordinate basis of the tangent space are Poisson commuting. In fact they are commuting Hamiltonians of Hitchin systems [25]. This property implies that the curvature of the connection is in fact a differential operator of second order rather than of third order, as one would a priori expect. Then one uses cohomological arguments to show that such that there are in fact no globally defined second order operators except for constants.

In our situation we also have the Poisson commutation of a principal part of the operators, as can be shown by r-martix techniques. In fact they describe the second order integrals of motion of Hitchin systems for curves with marked points. However the cohomological arguments are more tricky in this case.

A direct general proof, based on the classical Yang–Baxter equation, Theorem 4.5, would be instructive, and would have the advantage to give commutation of the differential operators ∇_j for general complex values of κ, and when acting on more general functions than conformal blocks, for which the differential operators can be considered.

In fact, Theorem 5.2 gives flatness in the directions of fixed complex structure in a universal form: it expresses the commutativity of the covariant derivatives, viewed as differential operators with coefficients in a tensor power of the universal enveloping algebra, without the need to mention representations.

In particular one could approach the interesting point $\kappa = 0$, which is related to a quantization of Hitchin systems, see [24, 25] and to the Beilinson–Drinfeld geometric Langlands correspondence, see [20].

It is likely that the proof of Theorem 5.2 can be applied to prove the flatness of the connection on the whole moduli space, but the technical details appear to be more involved.

One motivation for the construction described in this paper is the hope to understand the q-deformation of conformal field theory on Riemann surfaces. It turns out that in special cases the KZB equations admit a q-deformation, a system of compatible *difference* equations: see [21] and Varchenko's contribution to these proceedings for the genus zero case, and [14, 18], for the genus one case. In this paper, we have completed the first step in the "St. Petersburg q-deformation recipe" (see [11] and Faddeev's lectures in these proceedings): we have written all equations in terms of (a version of) classical r-matrices. The second step is to replace the classical r-matrices by quantum R-matrices, which for genera larger than one is an open problem.

Acknowledgments. I wish to thank K. Gawędzki and A. Varchenko for many useful discussions on this and related subjects, B. Enriquez for interesting correspondence and S. Kumar for his explanations. Also I am grateful to the organizers and the participants to the les Houches school for the pleasant and stimulating atmosphere. This research was partially supported by the NSF grant DMS-9400841.

References

[1] E. Arbarello, C. de Concini, V. G. Kac and C. Procesi, *Moduli spaces of curves and representation theory*, Commun. Math. Phys. **117** (1988), 1–36

[2] A. Beauville and Y. Laszlo, *Conformal blocks and generalized theta functions*, Commun. Math. Phys. **164** (1994), 335–340

[3] D. Bernard, *On the Wess–Zumino–Witten model on the torus*, Nucl. Phys. **B303** (1988), 77–93

[4] D. Bernard, *On the Wess–Zumino–Witten model on Riemann surfaces*, Nucl. Phys. **B309** (1988), 145–174

[5] D. Bernard and G. Felder, *Fock representations and BRST cohomology in SL(2) current algebra*, Commun. Math. Phys. **127** (1990), 145–168

[6] I. Cherednik, *Integral solutions of trigonometric Knizhnik–Zamolodchikov equations and Kac–Moody algebras*, Publ. RIMS **27** (1991), 727–744

[7] T. Eguchi and H. Ooguri, *Conformal and current algebras on a general Riemann surface*, Nucl. Phys. **B187** (1987) 127–134

[8] B. Enriquez and G. Felder, in preparation

[9] B. Enriquez and V. Rubtsov, *Hitchin systems, higher Gaudin operators and R-matrices*, Math. Res. Lett. **3** (1996), 343–357

[10] P. Etingof, I. Frenkel and A. Kirillov, Jr., *Spherical functions on affine Lie groups*, Duke Math. J. **80** (1995), 59–90

[11] L. D. Faddeev, L. A. Takhtajan, *Hamiltonian methods in the theory of solitons*, Springer, 1987

[12] F. Falceto and K. Gawędzki, *Chern–Simons states at genus one*, Commun. Math. Phys. **159** (1994), 549–579

[13] G. Faltings, *A proof for the Verlinde formula*, J. Alg. Geom. **3**(1994), 347–374

[14] G. Felder, *Conformal field theory and integrable systems associated to elliptic curves*, Proceedings of the International Congress of Mathematicians (Zürich 1994), p. 1247–1255, Birkhäuser, 1994

[15] G. Felder, K. Gawędzki and A. Kupiainen, *Spectra of Wess–Zumino–Witten models with arbitrary simple groups*, Commun. Math. Phys. **117** (1989), 117–158

[16] D. Friedan and S Shenker, *The analytic geometry of two-dimensional conformal field theory*, Nucl. Phys. **B281** (1987) 509–545.

[17] G. Felder and A. Varchenko, *Integral representation of solutions of the elliptic Knizhnik–Zamolodchikov–Bernard equations*, Int. Math. Res. Notices (1995), no. 5, 221–133

[18] G. Felder, V. Tarasov and A. Varchenko, *Solutions of the elliptic qKZB equations and Bethe ansatz I* , preprint 1996

[19] G. Felder and C. Wieczerkowski, *Conformal blocks on elliptic curves and the Knizhnik–Zamolodchikov–Bernard equation*, Commun. Math. Phys. **176** (1996), 133–162

[20] E. Frenkel, *Affine algebras, Langlands duality and Bethe ansatz*, Proceedings of the International Congress of Mathematical Physics (Paris 1994), 606–642, International Press, Boston, 1995

[21] I. Frenkel and N. Reshetikhin, *Quantum affine algebras and holonomic difference equations*, Commun. Math. Phys. **146** (1992), 1–60

[22] K. Gawędzki, SU(2) *WZW theory at higher genus*, Commun. Math. Phys. **169** (1995), 329–371

[23] P. A. Griffiths, *Topics in algebraic and analytic geometry*, Princeton University Press, Princeton, 1974

[24] N. Hitchin, *Flat connections and geometric quantization*, Commun. Math. Phys. **131** (1990), 347–380

[25] N. Hitchin, *Stable bundles and integrable systems*, Duke Math. J. **54** (1987), 91–114

[26] D. Ivanov, *Knizhnik–Zamolodchikov–Bernard equations on Riemann surfaces*, Int. J. Mod. Phys. **A10** (1995), 2507–2536

[27] V. G. Kac, *Infinite dimensional Lie algebras*, Cambridge University Press, 1990

[28] I. Krichever and S. Novikov, *Algebras of Virasoro type, the energy-momentum tensor, and operator expansions on Riemann surfaces*, Funct. Anal. Appl. **28** (1994), 21–32

[29] S. Kumar, M. S. Narasimhan and A. Ramanathan, *Infinite Grassmannians and moduli spaces of G-bundles*, Math. Ann. **300** (1994), 41–75.

[30] V. G. Knizhnik and A. B. Zamolodchikov, *Current algebra and Wess–Zumino model in two dimensions*, Nucl. Phys. **B247** (1984) 83–103.

[31] G. Pressley and G. Segal, *Loop groups*, Oxford University Press, 1986

[32] G. Segal, *Geometric aspects of quantum field theory*, Proceedings of the International Congress of Mathematicians (Kyoto 1990), 1387–1396, Math. Soc. Japan, Tokyo, 1991

[33] V. V. Schechtman and A. N. Varchenko, *Arrangements of hyperplanes and Lie algebra homology*, Inv. Math. **106** (1991), 139–194

[34] A. Tsuchiya, K. Ueno and Y. Yamada, *Conformal field theory on universal family of stable curves with gauge symmetry*, Adv. Stud. Pure Math., **19** (1989), 459–566

[35] A. N. Tyurin, *Periods of quadratic differentials*, Russ. Math. Surv. **33**:6 (1978), 169–221

[36] E. Witten, *Non-Abelian bosonization*, Commun. Math. Phys. **92** (1984), 455–472